

APPLICANT: Fischhoff, David A.
APPLICANT: Fuchs, Roy L.
APPLICANT: Lavrik, Paul B.
APPLICANT: McPherson, Sylvia A.
APPLICANT: Perlak, Frederick J.
TITLE OF INVENTION: Insect Resistant Plants
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lawrence M. Lavlin, Jr., Monsanto Co., BBAF
STREET: 700 Chesterfield Parkway No. 5495071th
CITY: St. Louis
STATE: Missouri
COUNTRY: United States of America
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,281
FILING DATE: 19930604
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/523284
FILING DATE: 14-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Lavlin Jr., Lawrence M.
REGISTRATION NUMBER: 30,768
REFERENCE/DOCKET NUMBER: 38-21(10629)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 537-7286
TELEFAX: (314) 537-6047
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2615 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 205..2139
US-08-072-281-1

Query Match 11.6%; Score 31.6; DB 1; Length 2615;
Best Local Similarity 51.4%; Pred. No. 1.4;
Matches 73; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 1 GTTGAATCGACCCCAATGTCGACAGATGAACACATCTCATTCACAAATAAGACATG 60
DB 2163 GTGACCATCTATGATAGTAAAGCAAGATGAATAAATGATTCATTAATAACATA 2222
QY 61 GTGACCCAGCATCTCTATATCTGTTGAAGCTACAGCAAGCATTTATTTCAAAAT 120
DB 2223 GTGTTCTTCAACTTCGCTTTTGAAGGTAGATGAAGAACTATTTTATTTCAAAAT 2282
QY 121 GTTGCATTTTATTCCTGAAA 142
DB 2283 GAAGGAAGTTTAAATATGTAA 2304

RESULT 3
US-08-759-446-1
Sequence 1, Application US/08759446
GENERAL INFORMATION:
APPLICANT: Fischhoff, David A.
APPLICANT: Fuchs, Roy L.
APPLICANT: Lavrik, Paul B.
APPLICANT: McPherson, Sylvia A.
APPLICANT: Perlak, Frederick J.

TITLE OF INVENTION: Insect Resistant Plants
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lawrence M. Lavlin, Jr., Monsanto Co., BBAF
STREET: 700 Chesterfield Parkway No. 5763241th
CITY: St. Louis
STATE: Missouri
COUNTRY: United States of America
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,446
FILING DATE: 05-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/072,281
FILING DATE:
FILING DATE: 14-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Lavlin Jr., Lawrence M.
REGISTRATION NUMBER: 30,768
REFERENCE/DOCKET NUMBER: 38-21(10629)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 537-6047
TELEFAX: (314) 537-7286
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2615 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 205..2139
US-08-759-446-1

Query Match 11.6%; Score 31.6; DB 1; Length 2615;
Best Local Similarity 51.4%; Pred. No. 1.4;
Matches 73; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 1 GTTGAATCGACCCCAATGTCGACAGATGAACACATCTCATTCACAAATAAGACATG 60
DB 2163 GTGACCATCTATGATAGTAAAGCAAGATGAATAAATGATTCATTAATAACATA 2222
QY 61 GTGACCCAGCATCTCTATATCTGTTGAAGCTACAGCAAGCATTTATTTCAAAAT 120
DB 2223 GTGTTCTTCAACTTCGCTTTTGAAGGTAGATGAAGAACTATTTTATTTCAAAAT 2282
QY 121 GTTGCATTTTATTCCTGAAA 142
DB 2283 GAAGGAAGTTTAAATATGTAA 2304

RESULT 4
US-09-027-998A-1
Sequence 1, Application US/09027998A
GENERAL INFORMATION:
APPLICANT: Fischhoff, David A.
APPLICANT: Fuchs, Roy L.
APPLICANT: Perlak, Frederick J.
TITLE OF INVENTION: Insect Resistant Plants
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold White and Durkee
STREET: PO Box 4433

1
 2 CITY: Houston
 3 STATE: TX
 4 COUNTRY: USA
 5 ZIP: 77210-4433
 6
 7 COMPUTER READABLE FORM:
 8 MEDIUM TYPE: Floppy disk
 9 COMPUTER: IBM PC compatible
 10 OPERATING SYSTEM: PC-DOS/MS-DOS
 11 SOFTWARE: PatentIn Release #1.0, V
 12 CURRENT APPLICATION DATA:
 13 APPLICATION NUMBER: US/09/027,998A
 14 FILING DATE: 23-FEB-1998
 15
 16 CLASSIFICATION:
 17 ATTORNEY/AGENT INFORMATION:
 18 NAME: Patterson, Melinda L
 19 REGISTRATION NUMBER: 33,062
 20 REFERENCE/DOCKET NUMBER: MOBJ:195
 21
 22 TELECOMMUNICATION INFORMATION:
 23 TELEPHONE: (713) 787-1400
 24
 25 INFORMATION FOR SEQ ID NO: 1:
 26 SEQUENCE CHARACTERISTICS:
 27 LENGTH: 2615 base pairs
 28 TYPE: nucleic acid
 29 STRANDEDNESS: double
 30 TOPOLOGY: Linear
 31
 32 US-09-027-998A-1

Query Match	11.68	Score 31.6	DB 4	Length 2615
Best Local Similarity	51.48	Pred. No. 1.4		
Matches 73	Conservative	0	Mismatches 69	Indels 0
			Gaps	0

Accession	Sequence	Position
QY	1 GTTGCACATCACCCCAATGTGCCAGGAAGATGAAACACCTATTCCAAACAAATTAAGACACATG	60
Db	2163 GTGACCACTATGATAGTAAAGCAAGATATAAAATGATTCATTAATGAATTAACATA	2222
QY	61 GTGACCCAGGACATCTCTGTATCTGTTTGAAGCTACAGCAAGAAACGATTTATTTCAAAAT	120
Db	2223 GTGTCTTCAACCTTGCGCTTTTGAAGCTAGATGAAGAAACACTATTTTATTTTCAAAAT	2282
QY	121 GTTGCACATTTTGATTCCTGTAAA	142
Db	2283 GAAAGAAAGTTTAAATATGTAA	2304

RESULT 5
US-09-031-442A-21
Sequence 21, Application US/09031442A
Patent No. 5955310
GENERAL INFORMATION:
APPLICANT: Widner, William
APPLICANT: Sloma, Alan
APPLICANT: Thomas, Michael D.
TITLE OF INVENTION: Methods for Producing A Polypeptide
TITLE OF INVENTION: In A Bacillus Cell
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 59553100 No. 5955310disk of No. 5955310th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,442A
FILING DATE: 26-FEB-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5455,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELETYPE: 212-878-9655

INFORMATION FOR SEQ ID NO: 21
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3050 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 OS-09-031-442A-21

Query Match	11.6%	Score 31.6	DB 2	Length 3050
Best Local Similarity	51.4%	Pred. No. 1.5	Mismatches 69	Indels 0
Matches 73	Conservative 0	Mismatches 69	Indels 0	Gaps 0
QY	1 GTTGCAATGACCCCATGTGCGCAGAGATGAACACATTCATTCACAACAATAAGACATG	60		
DB	2594 GTGACCATCTATGATGTAGTAAAGCAAAAGATAAAAAATGAGTTCATTAATAAGATACATA	2653		
QY	61 GTGACCCAGGACATCTCTGTATCTGTTTGAAGCTACAGAGAAAGCATTTTATTTCAAAAT	120		
DB	2634 GTGCTTCACACTTCGCTTTTGAAGCTAGATGAAGAACACTATTTTATTTTCAAAAT	2713		
QY	121 GTTGCAATTTTGATTCCTGAA	142		
DB	2714 GAAGGAAGTTTAAATATGTAA	2735		

RESULT: 6
 US-09-238-377-21
 Sequence 21, Application US/09258377
 Patent No. 6255076
 GENERAL INFORMATION:
 APPLICANT: Widner, William
 APPLICANT: Sloma, Alan
 APPLICANT: Thomas, Michael D.
 TITLE OF INVENTION: Methods for Producing A polypeptide In a
 TITLE OF INVENTION: Bacillus Cell
 FILE REFERENCE: 5455,200-US
 CURRENT APPLICATION NUMBER: US/09/258,377
 CURRENT FILING DATE: 1999-02-26
 EARLIER APPLICATION NUMBER: 09/031,442
 EARLIER FILING DATE: 1998-02-26
 NUMBER OF SEQ. ID NOS: 33
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ. ID NO 21
 LENGTH: 3050
 TYPE: DNA
 ORGANISM: a:Bacillus
 US-09-258-377-21

Query Match	11.6%	Score 31.6	DB 4	Length 3050
Best Local Similarity	51.4%	Pred. No. 1.5		
Matches 73	Conservative	0	Mismatches 69	Indels 0
			Gaps	0
QY 61	GTTCGACATGACCCCATGTGCCAGAAAGATGAACATCTATTCAACAATAAAGACATG	60		
DB 2534	gtagaccatcatgatatgataagaacaggataaaaaaatgagltcataaatagaataacata	2653		
QY 61	GTACACCAGACATCTGTATCTGTTTACGCTACAGGAAGACATTTTATTCCAAAAT	120		
DB 2634	gtagtcttcaacttctgccttttgaaggtagatgaagaacatactatttatttcaaat	2713		
QY 141	GTTCGCATTTTGATTCCTGAAT	142		
DB 2714	gaaggaagtlttcaataatagtaa	2735		

NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.0
SEQ ID NO 11
LENGTH: 3663
TYPE: DNA
ORGANISM: Homo sapiens
US-09-499-884-11

Query Match 11.1%; Score 30.4; DB 4; Length 3663;
Best Local Similarity 61.2%; Pred. No. 3.8;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

cy 18 TGTCCAGAGATGAAACACTCATTCACAATAAAGACATGGTGAACCGAGCATCTCT 77
| | | | | | | | | | | | | | | | | | | | |
Db 939 tgggcaacatgaagaactctgttctccaaaaataaaataaataggcagcgcatgtc 998

cy 78 GTATCTGTTTGAGCTCAG 97
||| ||| | |||
Db 999 gtagcaccttgattccag 1018

RESULT 9
US-08-642-846-1
Sequence 1, Application US/08642846
Patent No. 5886151

GENERAL INFORMATION:
APPLICANT: HOSTETTER, MARGARET K.
APPLICANT: GALE, CHERYL A.
APPLICANT: BENDEL, CATHERINE M.
APPLICANT: TAO, NIAN-JUN
TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: MUETING, RAASCH, GERHARDT & SCHWAPPACH, P.A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOCS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,846
FILING DATE: 03-MAY-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MUETING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110,00280101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5194 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-642-846-1

Query Match 10.8%; Score 29.6; DB 2; Length 5194;
Best Local Similarity 49.0%; Pred. No. 7.6;
Matches 77; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

58 ATGTGACCAGGCATCTCTGTATCTGTTTGAAGCTACAGAAAGCAATTTATTTCAAA 117

[illegible]

```

OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (3326)..(3327)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (3734)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (3888)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (4412)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (4607)..(4608)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (4656)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (4666)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (4736)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (4761)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
FEATURE:
NAME/KEY: unsure
LOCATION: (5354)
OTHER INFORMATION: n may be A or G or C or T/U, or the nucleotide is
OTHER INFORMATION: not present
US-09-289-040-1

Query_Match      10.8%; Score 29.4; DB 4; Length 5401;
Best Local Similarity 48.2%; Pred. No. 8.8;
Matches 107; Conservative 0; Mismatches 114; Indels 1; Gaps 1;

Gy 24 AGAAGATGAACACTCATTCATTCACAACAATAAAGACATGGTGAGCCAGGCATCTGTATCT 83
Db 35 AAAAGTTGATATTTATGAAAATTCATATTAAGACGAATCTAACMAAGATCCCATGAAATAT 326
Gy 34 GTTTGAACCTACAGAA-AGCGATTTTATTTCAAAAAAATGTTGCCATTTTGATTCCTGAAA 142
Db 35 GTTTTATATATAGTAATAAGAAATTTTGGTCMAAAATTTTAAATTTTGACTCCTAANA 266
Gy 133 CATGAACACAAAGAGNTGACTATGTGAGCCAAACTGAGACCTACACAAAATGCTGATG 202
Db 285 TAGAAAGGGGACTCATATTAATGAGACGATGGAATATTTGTTTACTAGTTTAGAAAG 206
Gy 203 TTTCGTGTCGAGTCTANTCCTCCAGNMAATGATGAACCT 244

```


Db 205 TATTAGATGCTTATACATAGCGATGAGGTCAAGACAATTTCT 164

RESULT 11

US-08-441-139-8

Sequence 8, Application US/08441139
Patent No. 5773245

GENERAL INFORMATION:

APPLICANT: Witterup, Dr. Karl D.

APPLICANT: Robinson, Anne S.

TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF

NUMBER OF SEQUENCES: 20

RECOMBINANTLY EXPRESSED PROTEINS

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: NY

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/441,139

FILING DATE: 15-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/089,997

FILING DATE: 06-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Digilio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 8646

TELECOMMUNICATION INFORMATION:

TELEPHONE: 516-742-4343

TELEFAX: 516-742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 6030 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1004..4753

US-08-441-139-8

Query Match 10.78; Score 29.2; DB 1; Length 6030;

Best Local Similarity 57.18; Pred. No. 11;

Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Db 126 CATTGTGATTCCTGAGACATGAGACAGAGAGGATGATGAGACCAAACTTGAGAC 185

Db 4603 CATTCTACTACTGACAAACAGAAATGAAATGATTAATACAGAAATACATGAATT 4662

QY 186 CTACAAAATGCTGATGTTCTGTTGCTGAG 216

Db 4663 GCAAAAGATTCTTATTCATTTGAGAGCTGAG 4693

RESULT 12

US-08-485-216-1

Sequence 1, Application US/08485216

Patent No. 5695960

GENERAL INFORMATION:

APPLICANT: Chan, Yoon Loong

APPLICANT: Hanl, Eric K.

TITLE OF INVENTION: HIPURICASE GENE

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merchant & Gould

STREET: 3100 No. 5695960west Center, 90 S. 7th Street

CITY: Minneapolis

STATE: Minnesota

COUNTRY: U.S.A.

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,216

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO 94/26907

FILING DATE: 13-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: Mueller, Douglas P.

REGISTRATION NUMBER: 30,300

REFERENCE/DOCKET NUMBER: 7933,290502

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612/332-5300

TELEFAX: 612/332-9081

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1338 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1338

US-08-485-216-1

Query Match 10.68; Score 29; DB 1; Length 1338;

Best Local Similarity 50.48; Pred. No. 7;

Matches 68; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Db 154 AAGGNTGACATGTCGAGACCAAACTTGAGACTACAAAAATGCTGATGTTGCT 213

Db 220 CAAGCGAATTTGAAAAATTCGTCATCAATTCATGAAATCTGAGCTTGTTGAT 279

QY 154 AAGGNTGACATGTCGAGACCAAACTTGAGACTACAAAAATGCTGATGTTGCT 213

Db 220 CAAGCGAATTTGAAAAATTCGTCATCAATTCATGAAATCTGAGCTTGTTGAT 279

QY 214 GAGTCTANTCTCCA 228

Db 230 GAATTATCTACTGCA 294

RESULT 13

US-09-003-245-1

Sequence 1, Application US/09003245

Patent No. 5981189

GENERAL INFORMATION:

APPLICANT: Chan, Yoon L.

APPLICANT: Hanl, Eric K.

TITLE OF INVENTION: Hippuricase Gene

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: MERCHANT & GOULD

STREET: 3100 No. 5981189west Center, 90 South Seventh Street

CITY: Minneapolis

STATE: Minnesota

COUNTRY: U.S.A.

ZIP: 55402-4131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/003,245
FILING DATE: 06-JAN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mueller, Douglas
REGISTRATION NUMBER: 30,300
REFERENCE/DOCKET NUMBER: MAG 7933.29-US-D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (612) 332-5300
TELEFAX: (612) 332-9081
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1338 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Campylobacter jejuni
STRAIN: TGH9011
INDIVIDUAL ISOLATE: ATCC43431
IMMEDIATE SOURCE:
LIBRARY: C. jejuni genomic library in pBR322
CLONE: PHIP-O
POSITION IN GENOME:
MAP POSITION: 380 kb SmaI fragment
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1338
US-09-003-245-1

Query Match 10.6%; Score 29; DB 2; Length 1338;
Best Local Similarity 50.4%; Pred. No. 7;
Matches 68; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 94 ACAGAGAAAGCGATTATTTCAAAAATGTTGCCATTGATTCCTGAAACATGGAAGACA 153
DB 160 AGAGAAATTTATACAGAGAAATAAGATTGAATTTATTCAGAAATACAGACTTA 219
QY 154 AAGGTGACTATGTGAGACCAAACTTGAGACTTACAAAAAGCTGATGTTCTGGTCT 213
DB 220 CAAGCGCAATTTGAAAAAATTCGTCATCAAAATCGAAAAATCCGAGCTGGTTTGAT 279
QY 214 GAGTCTANTCTCCA 228
DB 280 GAATTAATGACTGCA 294

RESULT 14
US-08-853-552-1
Sequence 1, Application US/08853552
Patent No. 6013501
GENERAL INFORMATION:
APPLICANT: Chan, Yoon Loong
APPLICANT: Hanl, Eric K.
TITLE OF INVENTION: HIPURICASE GENE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 3100 No. 6013501west Center, 90 S. 7th Street
CITY: Minneapolis
STATE: Minnesota
COUNTRY: U.S.A.
ZIP: 55402
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,552
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: 08/485,216
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mueller, Douglas P.
REGISTRATION NUMBER: 30,300
REFERENCE/DOCKET NUMBER: 7933.29US02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/332-5300
TELEFAX: 612/332-9081
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1338 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1338
US-08-853-552-1

Query Match 10.6%; Score 29; DB 3; Length 1338;
Best Local Similarity 50.4%; Pred. No. 7;
Matches 68; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 94 ACAGAGAAAGCGATTATTTCAAAAATGTTGCCATTGATTCCTGAAACATGGAAGACA 153
DB 160 AGAGAAATTTATACAGAGAAATAAGATTGAATTTATTCAGAAATACAGACTTA 219
QY 154 AAGGTGACTATGTGAGACCAAACTTGAGACTTACAAAAAGCTGATGTTCTGGTCT 213
DB 220 CAAGCGCAATTTGAAAAAATTCGTCATCAAAATCGAAAAATCCGAGCTGGTTTGAT 279
QY 214 GAGTCTANTCTCCA 228
DB 250 GAATTAATGACTGCA 294

RESULT 15
US-08-685-625A-5/C
Sequence 5, Application US/08685625A
Patent No. 5945301
GENERAL INFORMATION:
APPLICANT: UENO, Naoto
APPLICANT: MATSUMOTO, Kunihiko
APPLICANT: IRIE, Kenji
TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,625A

FILING DATE: 24-JUL-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-253549
 FILING DATE: 29-SEP-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Meuth, Donna M.
 REGISTRATION NUMBER: 36,607
 REFERENCE/DOCKET NUMBER: 001560-267
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2656 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 183..1922
 US-08-685-625A-5

Query Match 10.6% Score 29; DB 2; Length 2656;
 Best Local Similarity 53.2% Pred. No. 9;
 Matches 84; Conservative 0; Mismatches 71; Indels 3; Gaps 1;
 OY 45 ACNAAATAAGGACATGTGACCCAGCATCTCTGTATCTGTTGAAGCT--ACAGGAAA 101
 DB 2209 ACAATGACATGCATATACAGGACAGTTCATCTGCTGTTGAAGTGTGCAAGTGCT 2150
 OY 102 GCGATTTTATTCAAAAATGTTGCCATTTTGATCTCTGAACATGGAAGACAAAGNTGA 161
 DB 2149 GCTCATTCAGTCACAGCATGTATCCATGTATGCAATGAAACAGTAAATGTATGTCCA 2090
 OY 162 CTATGTGAGACCAAACTGAGACCTACAAAATGCTG 199
 DB 2089 CCATGAGAAAGAAATTAACATGAAAAAATGCGAG 2052

Search completed: April 3, 2002, 20:36:15
 Job time: 40733 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 02:55:32 ; Search time 22700.8 Seconds
(Without alignments)
198.395 Million cell updates/sec

Title: US-09-049-696-2

Perfect score: 273
Sequence: 1 GTTCAATGACCCCAATGT.....AGATGGCACTGTGGCGAG 273

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_ov:*
21: em_or:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_hqo_hum:*
31: em_hqo_inv:*
32: em_hqo_rnd:*
33: em_hqo_hum:*
34: em_hqo_inv:*
35: em_hqo_rnd:*
36: em_hqo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	267.4	97.9	2826	9 AF127036	AF127036 Homo sapi
2	267.4	97.9	3311	6 AX193489	AX193489 Sequence
3	267.4	97.9	3311	6 AF039400	AF039400 Homo sapi
4	245.4	89.9	401	6 AX192467	AX192467 Sequence
5	206.2	75.5	3079	4 AF095584	AF095584 Sus scrofa
6	182.2	66.7	2937	10 AB017156	AB017156 Mus muscu
7	151.2	55.4	3204	9 AF127035	AF127035 Homo sapi
8	151.2	55.4	3221	9 AK000072	AK000072 Homo sapi
9	151.2	55.4	3265	6 AF092338	AF092338 Sequence
10	143	52.4	35278	6 AF039401	AF039401 Homo sapi
11	143	52.4	164891	2 AL358950	AL358950 Homo sapi
12	141.4	51.8	113764	9 HSJ651E0	AL122002 Human DNA
13	124.6	45.6	1265	4 AF001264	AF001264 Bos tauru
14	124.6	45.6	2820	4 AF001263	AF001263 Bos tauru
15	124.6	45.6	3288	4 AF001262	AF001262 Bos tauru
16	124.6	45.6	3317	4 AF001261	AF001261 Bos tauru
17	123	45.1	2984	4 BTU36445	U36445 Bos tauru
18	119.8	43.9	3415	9 AF043976	AF043976 Homo sapi
19	115	42.1	2765	10 AF115852	AF115852 Mus muscu
20	115	42.1	3058	10 AF108501	AF108501 Mus muscu
21	115	42.1	3137	10 BC008147	BC008147 Mus muscu
22	113.4	41.5	3022	10 AF047838	AF047838 Mus muscu
23	113.4	41.5	3471	10 AF052746	AF052746 Mus muscu
24	108	39.6	375	6 AX071171	AX071171 Sequence
25	75.2	27.5	2832	9 AF127980	AF127980 Homo sapi
26	75.2	27.5	2970	9 AF043977	AF043977 Homo sapi
27	75.2	27.5	4077	6 AX054697	AX054697 Sequence
28	73.6	27.0	3604	9 AB026833	AB026833 Homo sapi
29	59.8	21.9	175591	2 AC068071	AC068071 Homo sapi
30	55	20.1	124096	2 AL354988	AL354988 Homo sapi
31	48.6	17.8	227	6 AX186451	AX186451 Sequence
32	48.6	17.8	350	6 AX188011	AX188011 Sequence
33	48.6	17.8	454	6 AX185116	AX185116 Sequence
34	42.4	15.5	140718	2 AL356270	AL356270 Homo sapi
35	39.8	14.6	65646	9 AL512325	AL512325 Human DNA
36	39	14.3	146028	2 AC021573	AC021573 Homo sapi
37	38	13.9	1315	8 HNY247165	AL247165 Henna
38	38	13.9	163027	2 AC022846	AC022846 Homo sapi
39	38	13.9	175590	2 AF307159	AF307159 Homo sapi
40	38	13.9	181848	2 AC090199	AC090199 Homo sapi
41	37.4	13.7	60791	9 AC004749	AC004749 Homo sapi
42	37.4	13.7	79041	9 AC009016	AC009016 Homo sapi
43	37.4	13.7	79851	9 AC034252	AC034252 Homo sapi
44	37.4	13.7	122557	2 AC011429	AC011429 Homo sapi
45	37.4	13.7	131529	2 AC008497	AC008497 Homo sapi

ALIGNMENTS

RESULT: 1
AF127036
LOCUS: 1
DEFINITION: Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)
ACCESSION: AF127036
VERSION: AF127036
KEYWORDS: complete cds.
SOURCE: human
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE: 1 (bases 1 to 2826)
AUTHORS: Agnel, M., Vermet, T., and Culouscou, J.M.
TITLE: Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea
JOURNAL: FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE: 99364503
PUBMED: 10437792

REFERENCE 2 (bases 1 to 2826)
 AUTHORS Agnel, M. and Culouscou, J.-M.
 TITLE Direct Submission
 JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHLABO, 10 Rue des Carrières, Rueil-Malmaison 92500, France

FEATURES
 Source Location/Qualifiers
 1..2826

gene /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue-type="small intestine; colon"
 1..2826
 /gene="CACCL1"
 5..2749
 /gene="CACCL1"
 /note="bovine epithelial chloride channel homolog"
 /codon_start=1
 /product="calcium-activated chloride channel protein 1"
 /protein_id="A025487.1"
 /db_xref="GI:4583469"
 /translation="MGPKSSVFFILHLLEGALSNSLIQLNNNGEIVVAIDPNP
 EDLELIQIKDMVTOASLYLEFATGKREYFNKVALIPEWTKTKADYVRPKLETKNA
 DIVLVESTPGNDDEPTETOMGCKEGERIHLPDPIAGKIAEYGPGRAPVHMAH
 LRNGVPEYNDDEKFLNSGRIOAVKCSAGITGVNKKCGSCYTRCFENKVTGL
 YKGEFEVLSRQTEKASIMFAQHVDSIVECTEQNHKKEAPNKNOCNRSSTWEVI
 RDSSEFKTTPMTTPPNTFSLQIGRIYCLVDKSGMATGRNRLNDAQGLFL
 LQVELGSMVAGVTFDSAAHVSELIQINSQSDRDTLAKRLPAAAGSTISCSGLSA
 FTVIRKPTDGESEIVLTLDGENDTISGCFNVSQSGAIHTVALGPSAOLLEELSK
 MTGLOVTAASDOVONNGILDAFGALSGNSGVSORSIOLEKGLILOSOMNGVIV
 DSTVGGKDLFLITWTPROPOLLMDPSOKOGKGPVVDNTEKAYIQITGIAKVGWKY
 SIQASSQTLTLTVSRASNAFLPITVTSKTKNDTSKPSPLVYVANTROGAPILRA
 SVTALIESNKVTVLELDNGAGADAKDGVDSRYTDTYDNRYSVKYKALGVN
 AARRVYIPQSGALYIPGMIENDEIQNMPPEIKNDVOKHGVCFSTSSGSVAS
 DVNPAPIPDLEPPGQITDLKAEIHGSLINLTWPAFGDYDGHGTHKXIIIRSTSLD
 LRDFNESLQVNTALIRKEANSEVEFLFKPENTPEFGTDLFLALQAVDVYDKSEI
 SNIAVSLFIPQTPPEPPEPSPDETSAPCPNHINSTINGIHILKIMKMWIELQSLIA

BASE COUNT 875 a 623 c 632 g 696 t
 ORIGIN

Query Match 97.9%; Score 267.4; DB 9; Length 2826;
 Best Local Similarity 98.2%; Pred. No. 6.4e-64;
 Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTTGAATCGACCCCAATGTGCCAGAGATGAACACATCTATTCACAAATAAGACATG 60
 DB 113 GTTGAATCGACCCCAATGTGCCAGAGATGAACACATCTATTCACAAATAAGACATG 172
 QY 61 GTGACCCGAGCATCTCTGATCTGTTGAAGTACAGAAAGCATTTATTTCAAAAT 120
 DB 173 GTGACCCGAGCATCTCTGATCTGTTGAAGTACAGAAAGCATTTATTTCAAAAT 232
 QY 121 GTTGCATTGATTCCTGAAGAATGAGAAAGATGACTATGTGAGACCAAACTT 180
 DB 233 GTTGCATTGATTCCTGAAGAATGAGAAAGATGACTATGTGAGACCAAACTT 292
 QY 181 GAGACCTCAAAAATGCTGATCTTGTGCTGAGTCTANTCTCCAGGNAATGATGAA 240
 DB 293 GAGACCTCAAAAATGCTGATCTTGTGCTGAGTCTACTCTCCAGGNAATGATGAA 352
 QY 241 CCTTACTGNGCAGATGGGCAACGTGGCGAG 273
 DB 353 CCTTACTGNGCAGATGGGCAACGTGGCGAG 385

RESULT 2
 AX193489 3311 bp DNA PAT 15-AUG-2001
 LOCUS AX193489
 DEFINITION Sequence 1056 from Patent WO0149716.
 ACCESSION AX193489
 VERSION AX193489.1 GI:15211440
 KEYWORDS human.
 SOURCE

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3311)
 AUTHORS Xu, J., Lodes, M. J., Secrist, H., Benson, D. R., Meagher, M. J.,
 Stolk, J. A., King, G. E., Wang, T. and Jiang, Y.
 TITLE Compounds for immunotherapy and diagnosis of colon cancer and
 methods for their use
 JOURNAL Patent: WO 0149716-A 1056 12-JUL-2001;
 CORIXA CORPORATION (US)

FEATURES
 Source Location/Qualifiers
 1..3311
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 1028 a 692 c 742 g 849 t
 ORIGIN

Query Match 97.9%; Score 267.4; DB 6; Length 3311;
 Best Local Similarity 98.2%; Pred. No. 6.5e-64;
 Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTTGAATCGACCCCAATGTGCCAGAGATGAACACATCTATTCACAAATAAGACATG 60
 DB 450 GTTGAATCGACCCCAATGTGCCAGAGATGAACACATCTATTCACAAATAAGACATG 519
 QY 61 GTGACCCGAGCATCTCTGATCTGTTGAAGTACAGAAAGCATTTATTTCAAAAT 120
 DB 530 GTGACCCGAGCATCTCTGATCTGTTGAAGTACAGAAAGCATTTATTTCAAAAT 579
 QY 121 GTTGCATTGATTCCTGAAGAATGAGAAAGATGACTATGTGAGACCAAACTT 180
 DB 580 GTTGCATTGATTCCTGAAGAATGAGAAAGATGACTATGTGAGACCAAACTT 639
 QY 181 GAGACCTCAAAAATGCTGATCTTGTGCTGAGTCTANTCTCCAGGNAATGATGAA 240
 DB 640 GAGACCTCAAAAATGCTGATCTTGTGCTGAGTCTACTCTCCAGGNAATGATGAA 699
 QY 241 CCTTACTGNGCAGATGGGCAACGTGGCGAG 273
 DB 700 CCTTACTGNGCAGATGGGCAACGTGGCGAG 732

RESULT 3

AF039400 3311 bp mRNA PRI 14-DEC-1998
 LOCUS AF039400
 DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA,
 complete cds.
 ACCESSION AF039400
 VERSION AF039400.1 GI:4009457
 KEYWORDS

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3311)
 AUTHORS Gruber, A. D., Elble, R. C., Ji, H. L., Schreier, K. D., Fuller, C. M. and
 Pauli, B. U.
 TITLE Genomic cloning, molecular characterization, and functional
 analysis of human CLCA1, the first human member of the family of
 Ca²⁺-activated Cl⁻ channel proteins
 JOURNAL Genomics 54 (2), 200-214 (1998)

REFERENCE 2 (bases 1 to 3311)
 MEDLINE Gruber, A. D., Elble, R. and Pauli, B. U.
 JOURNAL Direct Submission
 TITLE Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
 Cornell University, Ithaca, NY 14853, USA

FEATURES
 Source Location/Qualifiers
 1..3311
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"

gene, /map="1p22-p31"
/tissue_type="small intestine"
1. .3311
/gene="hclcal1"
352. .3096
/gene="hclcal1"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/db_xref="GI:4009458"

/translation="MGRFSSVFLILHLLEGALSLSLQLNNGTEGIVADNPV
EDLILQIKDVTQASLYLFATGKRFKFNVALIPETWTKADYVPKLETKYNA
DLVAEPPGDEPTTEQMGNGEKGERIHLTPDIACKLAETGPOKAFHEWVAH
LKMGEVDEPNDEKPYLSNGRIQAVRCSTGNTVKKCCGSCCTKCTEKVGL
YKGCSEVLSQSTOKESIMFAOHVDSIYEFCEQHNKAPKOKNOKNLSRTWEL
RSDPEKTTPTMOTPNPTPSILOIGORIVCLVLDKSGMATNGNLNNOGLFL
LOTVELGSVGVTPDSNAHVSELIQINSGRDRLARLPAAAGSTIGCGLSA
FTVIRKTPDSEELVLTGDEDTISGCFNEVKSGALIHVALPSAQLEELISK
MTGQIYASDVQNNGLIDAFGALSNGAVSORSIOLESKGLTQNSQNMNGIVY
DSTGKDLFLITWTQPPQILMPSCQKGGFYVDKTKMAVLOIPILAKGTRY
SLAASQTLFLVTSRASNATLPITVTSKTKNDTSKPPSLVYVANIPOGASPIRA
SVTALIESVNGKTVLELIDNGAGADATDDGVSYFTTYDNGRYSVKVRLGIVN
AARVYIPQSSALYIPWIEDELOMNPPEINQDDVQHKVCSTRISGGSPVAS
DVNPAPIDLPFPGQITDLKAEIHGSLNLNTWTAGDDYDGHGTAHRIIRKSTILD
LNDKNEISLQVNTALIPKEANSEVEFLKPNITFENGTDLFLAIOAVDKLKEI
SNIRVSLFIPPTPEPTSPDTSAPCNHINSYIPGHIILKIMMKWIGELQLSIA

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 97.9%; Score 267.4; DB 9; Length 3311;
Best Local Similarity 98.2%; Pred. No. 6.5e-64;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 GTTGCAATGACCCCAATGTCAGAGAAAGATGAACACTATTCACCAATTAAGAGACATG 60
|||||
DB 460 GTTGCAATGACCCCAATGTCAGAGAAAGATGAACACTATTCACCAATTAAGAGACATG 519
OY 61 GTGACCCAGGCAATCTGTATCTGTTGAAGCTACAGAGAAAGGATTTTATTTCAAAAT 120
|||||
DB 520 GTGACCCAGGCAATCTGTATCTGTTGAAGCTACAGAGAAAGGATTTTATTTCAAAAT 579
OY 121 GTTGCAATTTGATCTCTGTAACATGGAAGCAAGGNTGATCTATGTGAGACCAAACTT 180
|||||
DB 580 GTTGCAATTTGATCTCTGTAACATGGAAGCAAGGNTGATCTATGTGAGACCAAACTT 639
OY 181 GAGACCTACAAAATGCTGATGTTGCTGAGTCTANTCTCCAGGNAATGATGAA 240
|||||
DB 640 GAGACCTACAAAATGCTGATGTTGCTGAGTCTACTCTCCAGGNAATGATGAA 699
OY 241 CCTACACTGNGCAGATGGGCAACTGTGGGAG 273
|||||
DB 700 CCTACACTGNGCAGATGGGCAACTGTGGGAG 732

RESULT 4
AX192467 401 bp DNA PAT 15-AUG-2001
LOCUS AX192467
DEFINITION Sequence 34 from Patent WO0149716.
ACCESSION AX192467
VERSION AX192467.1 GI:15210431
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 401)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use

JOURNAL Patent: WO 0149716-A 34 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES location/Qualifiers
source 1. .401
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 127 a 80 c 95 g 99 t
ORIGIN

Query Match 89.9%; Score 245.4; DB 6; Length 401;
Best Local Similarity 97.5%; Pred. No. 6.6e-58;
Matches 268; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

OY 1 GTTGCAATTCACCCCAATGTCAGAGAAAGATGAACACTATTCACCAATTAAGAGACATG 60
|||||
DB 125 GTTGCAATTCACCCCAATGTCAGAGAAAGATGAACACTATTCACCAATTAAGAGACATG 84
OY 31 GTGACCCAGGCAATCTGTATCTGTTGAAGCTACAGAGAAAGGATTTTATTTCAAAAT 120
|||||
DB 85 GTGACCCAGGCAATCTGTATCTGTTGAAGCTACAGAGAAAGGATTTTATTTCAAAAT 144
OY 121 GTTGCAATTTGATCTCTGTAACATGGAAGCAAGGNTGATCTATGTGAGACCAAACTT 180
|||||
DB 145 GTTGCAATTTGATCTCTGTAACATGGAAGCAAGGNTGATCTATGTGAGACCAAACTT 204
OY 181 GAGACCTACAAAATGCTGATGTTGCTGAGTCTANTCTCCAGGNAATGATGAA 239
|||||
DB 205 GAGACCTACAAAATGCTGATGTTGCTGAGTCTACTCTCCAGGNAATGATGAA 264
OY 240 ACCCTACACTGNGCAGAT-GGGCAACTGTGGGAG 273
|||||
DB 265 ACCCTACACTGNGCAGATGGGCAACTGTGGGAG 299

RESULT 5
AF095584 3079 bp mRNA MM 13-OCT-2000
LOCUS AF095584
DEFINITION Sus scrofa epithelial chloride channel protein (AEC) mRNA,
complete cds.
ACCESSION AF095584
VERSION AF095584.1 GI:6002645
KEYWORDS
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 3079)
AUTHORS Gaspar,R.J., Racette,K.J., Gordon,J.R., Loewen,M.E. and
Forsyth,G.W.
TITLE Cloning a chloride conductance mediator from the apical membrane of
porcine ileal enterocytes
JOURNAL Physiol. Genomics (Online) 3 (2), 101-111 (2000)
MEDLINE 20473747
PUBMED 11015605
REFERENCE 2 (bases 1 to 3079)
AUTHORS Gaspar,R.J., Gabriel,S.E., Racette,K.J. and Forsyth,G.W.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-1998) Veterinary Physiological Sciences,
University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4,
Canada

FEATURES
source 1. .3079
location/Qualifiers
1. .3079
/organism="Sus scrofa"
/db_xref="taxon:9823"
/tissue_type="ileal mucosa"
1. .3079
/gene="AEC"
133. .2886
/gene="AEC"
/function="Involved in exocrine chloride conductance"
/codon_start=1
/product="epithelial chloride channel protein"

FEATURES	source
JOURNAL	Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea
MEDLINE	99364503
PUBMED	10437792
REFERENCE	2 (bases 1 to 3204)
AUTHORS	Agnel, M. and Culouscou, J.-M.
TITLE	Direct Submission
JOURNAL	Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrieres, Rueil-Malmaison 92500, France
location/Qualifiers	
1. .3204	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/tissue_type="colon"	
1. .3204	
/gene="CaCC2"	
29. .2782	
/note="bovine epithelial chloride channel homolog"	
/cdon_start=1	
/product="calcium-activated chloride channel protein 2"	
/protein_id="AA048398.1"	
/db_xref="GI:5726289"	
/translation="MGLRGEFVFLVLCILHOSNTSIFIKLNNGFEDIVIDPSVEE DEKIEIDEMVTASTYATFATEKEKREFFNVSILILENNKKNQYRPRKHENKHA VIAAPYLPGRDEPYTKQTECGEKEGEIHTPLLGLKKNQYRPGKLFVHEMALH RMVGFEDYNDDPEYTRAKSKKIEKTRCSAGISGNRRYKCOGGSCLSRACRIDSTLTL KCGDCEFPDKVOTERRASIMPMOSIDSVRECNKTNQERASLQNKCNRSWEYI SNSEDRKNTIPMTPPPPVPSLTKRORIVCLVLDSSGSGKDRLNRMOAKHFL LQTVENGSWGVHFEVDSATVTKLIDIKSDERNITLMAGLPYPLGGTSISGKIYA FOVIGELHSQDSEVLLDGEDNTASCSIDEVOKSGAIVHFTALGRADNAVIEMS KIVGSHFYVSDAQNGLIDAFGALTSYMTDLSOKSLQESKGLTNSAMNMDTVI IDSVGKDEFTLITMNSLSPSISLMDPSGTIMENFTIVATSKMAYLISPGAKYGVV YNIQAKNPELTITVTYSRANSVSPRTIVYNAKMKKNVSPSPMUYAELQGVV LGNVTAFTIESQNHTEVLELDNGACADSFKNQGVSRFTATTEGRSLAKRANG GANTRAKLRLPLNRAVITPGWVNGEITANPPPELDEDTQTTLEFSPRASGAVV VSVQSLPLDPQPPQITDIDATVHEDKILITLAPEDNFDVQKVRYSIRISASIL DLNDFEDALQVNTDLSPEANSKESFAPKEPNISENAITHIPIAKISDSMLTSK VSNIAQVTLFIQANPDIDIPPTPTPTPKSHNSGVNISTVLVIGSVIYVNFILS TTL"	

LOCUS	AK000072	3221 bp	mRNA	PRI	22-FEB-2000
DEFINITION	Homo sapiens cDNA FLJ20065 fis, clone COL01613, highly similar to ECLC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.				
ACCESSION	AK000072				
VERSION	AK000072.1	GI:7019922			
KEYWORDS	oligo capping; fis (full insert sequence).				
SOURCE	Homo sapiens colon cDNA to mRNA, clone_lib:COL clone:COL01613.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shihabara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.				
TITLE	NEBO human cDNA sequencing project				
REFERENCE	1 (sites)				
REFERENCE	2 (bases 1 to 3221)				
AUTHORS	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shihabara,T., Tanaka,T. and Nakamura,Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)				
COMMENT	NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'-6' end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).				
FEATURES	Location/Qualifiers				
source	1..3221				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="COL01613"				
	/clone_lib="COL"				
	/issue_type="colon"				
	/note="cloning vector pME18FL3"				
misc_feature	1..3221				
	/note="highly similar to ECLC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"				
BASE COUNT	1105 a 600 c 634 g 882 t				
ORIGIN					
Query Match	55.4%: Score 151.2; DB 9; Length 3221;				
Best Local Similarity	71.7%: Pred. No. 1.2e-31;				
Matches 195; Conservative	0; Mismatches 77; Indels 0; Gaps 0;				
1	TTGCAATGGACCCCAATGTGCGAAGATGGAACGCTCATCAACAATTAAGSCATGG	61			
2	TTGTTATAGATCCCTTAGTGTTCGACAGATGAAAAATATTGAACAATAATAGCGATATGG	208			
3	TGACCCAGCATCTCTCTATCTGTTGAAGCTACAGAAAGCATTTATTTCACAAATG	121			
4	TGACAGACGCTTCTACGACTGTTTGAAGCAGAAAGAAAGATTTTTCACAAATG	268			
5	TTGCCATTTTGATTCCTGAACATGGAAGACAAGAGTACTAGTGTGAGACCAAAACTTG	181			
6	TATCATATTATTCCTGATGGAATGGAGAAATCCCTCAGTACAAAGGCCAAATCATG	328			
7	AGACCTACAAAATGCTGATGTTCGTTGGCTGAGTCTAATTCCTCCAGGNAATGATGAC	241			
8	AAAACATTAACATGCTGATGTATTAGTTGACCAACCTTACACTCCACAGTAGAGATGAAC	388			
9	CTACACTGCGCAGATGGGCAACTGTGGCGAG	273			
10	CATACCAACAGCATTCACAGATGTGAGAG	420			
RESULT	9				

AX092338
LOCUS AX092338 3265 bp DNA
DEFINITION Sequence 69 from Patent WO0116318.
ACCESSION AX092338
VERSION AX092338.1 GI:13444483
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 3265)
AUTHORS Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gunney,A.U., Watanabe,C.K. and
Wood,W.I.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0116318-A 69 08-MAR-2001;
Genentech, Inc. (US)
FEATURES
source location/Qualifiers
1..3265
BASE COUNT 1159 a 596 c 632 g 878 t
ORIGIN
Query Match 55.4%; Score 151.2; DB 6; Length 3265;
Best Local Similarity 71.7%; Pred. No. 1.2e-31;
Matches 195; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
OY 2 TTGAATGACCCCAATGTCGACCAAGATGAACATCTATTCACCAATAAAGCATGG 61
Db 131 TTGTTATAGATCCTAGTGTGCGAAGATGAATAATTAATGACCAATAGAGATATG 190
OY 62 TGACCAAGCATCTCTATCTCTTTGAAGCTACAGAACGATTTATTTCAAAAATG 121
Db 191 TGACTACAGCTTACGACTGTTTGAAGCCACAGAAAAGATTTTTCAAAAATG 250
OY 122 TTGCAATTTGATTCCTGAACATGAAACAGCAAGATGACTATGTAGACCAAAATG 181
Db 251 TATCTAATTAATTCCTGAGAAATGGAAGAAATCCCTCACTACAAAAGCCAAACATG 310
OY 182 AGACCTCAAAAATGCGATGTTCTGTGCTGATCTATCTCCCGAGNAATATGATAC 241
Db 311 AAAACCTAAACATGCGATGTTATGATGACCACTACCTCCAGTAGAGATGAAC 370
OY 242 CCACTGNGCAGATGGGCAACTGTGGCGAG 273
Db 371 CATACACCAAGCACTTCACAGAAATGTGAGAG 402
RESULT 10
AF039401
LOCUS AF039401 35278 bp DNA
DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene,
complete cds.
ACCESSION AF039401
VERSION AF039401.1 GI:4009459
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 35278)
AUTHORS Gruber,A.D., Elble,R.C., Ji,H.U., Schreuer,K.D., Fuller,C.M. and
Pauli,B.U.
TITLE Genomic cloning, molecular characterization, and functional
analysis of human CLCA1, the first human member of the family of
Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 35278)
AUTHORS Gruber,A.D., Elble,R. and Pauli,B.U.

TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Department of Pathology, College of
Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES
source location/Qualifiers
1..35278
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
1..1617
/gene="hCLCA1"
1..33522
/gene="hCLCA1"
767..1265
repeat_region
/rpt_family="L1 transposable element"
/rpt_type="dispersed"
1587..1596
/gene="hCLCA1"
/note="putative"
1618..1708
/gene="hCLCA1"
/number=1
join(1618..1708,1962..2383,6667..6807,6979..7126,9696..9801,
9696..9801,15454..15631,18591..18809,19783..20010,22253..22427,
22253..22427,24522..24628,26644..26859,27447..27708,
28766..28936,31840..32079,32919..33522)
/gene="hCLCA1"
/product="calcium-dependent chloride channel-1"
1962..2383
/gene="hCLCA1"
/number=2
join(2222..2383,6667..6807,6979..7126,9696..9801,
15454..15631,18591..18809,19783..20010,22253..22427,
24522..24628,26644..26859,27447..27708,28766..28936,
31840..32079,32919..33522)
/gene="hCLCA1"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95429.1"
/db_xref="GI:4009460"
/translation="MGPRKSSVFLILHLLEGALSLSLIOLNNNGYEGIVAIIDPNP
EDFTLIQIKDMYQASLYLEAYGRFRFKNVALLIPETWKTKADYVRKLETYRNA
DVLVASTPEGNDPEYTEMNGCEKGERIHLPPDPIAGKRLAEYPOGAEYENAH
LRWGVDEYNNDEKFLSNGRIOAVCSAGITGTNVVKKCOGSGCYTRKCTFVGL
YERKGEFFVQSROTERKASIMFAOHVDSIYFCTEONHNKNEAPKKOKKCLRSTWVI
RDESDPKTKTPMTPDPNPFFSLIOIGORTVCLYLKSGSMATGRNLNRNQGQLEPL
LQVLELGSWGWTPRPSAAHVOSELQINSGRDYLARKLPAAAGSTVSGELKSA
FTYIRKRPYDGESEIVLTDGEDNTISGCFNEVKSQSAIILHVALPSAQLLELSK
MTGGLOTVASDOVQNGLLIDAFGLSSGNGVNSQRIQLESKGLTQNSQWNGTVIV
DSYVGRDTEFLITMTQPPQIILMDPSGQGGFVVDKNTKNAYLQIPIGAKGTWKY
SLQASSQFTLIVTSRASNAIPLPIVTSKTKDTSKPSPLVAVVNIROGASPIIRA
SVTALIESVNGKIVTLELDNGAGADATDDGVSYSTRYTDYDNGSYSVKRVGCVN
AARRYVPOOSGALYIPGILNDELQMNPRPRINDDYDQHKVCSTRISGGSFVAS
DYFNAPIDLPFGQITDLNAETHGSLNLMTWPAAGDYDHGTAKYIIRISTSLD
LRKNEVSLQVNTALIPKEANSEEVFLKPEPNTENQTDLEFIALQAVDKLKEEI
SNARVSLFIPQTPPETSPDETSPACPNHINSTRIPGIIHLIKIMKWLIGELQSLA
6667..6807
/gene="hCLCA1"
/number=3
6979..7126
/gene="hCLCA1"
/number=4
9696..9801
/gene="hCLCA1"
/number=5
15454..15631
/gene="hCLCA1"
/number=6
18591..18809
/gene="hCLCA1"
/number=7
19783..20010

Query Match	52.4%	Score 143;	DB 9;	Length 35278;
Best Local Similarity	99.3%	Pred. No. 2.8e-29;		
Matches 143; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

RESULT	11
AL358950/c	
LOCUS	
DEFINITION	Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING INFORMATION ***
DESCRIPTION	16481 bp DNA HTG 23-JUN-2001
PROGRESS	***, 10 unordered pieces.

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 164891)
Plumb, B.
Direct Submission
Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,

Center: Sanger Centre

Web site: <http://www.sanger.ac.uk>

----- Project Information
Contract number: 1744010

----- Summary Statistics -----

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least Q40
Consensus quality: 162610 bases at least Q30
Consensus quality: 163338 bases at least Q20
Insert size: 163991; sum-of-contigs
Insert size: 163373; 6.6% error; agarose-*fp*
Quality coverage: 5.19x in Q20 bases; sum-of-contigs quality
coverage: 5.21x in Q20 bases; agarose-*fp*

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	34236:	contly of	34236	bp	in length
*	1	34337	34336:	gap of	100	bp
*	34337	47372:	contly of	13036	bp	in length
*	47373	47472:	gap of	100	bp	
*	47473	67025:	contly of	19553	bp	in length
*	67026	67125:	gap of	100	bp	
*	67126	71233:	contly of	4114	bp	in length
*	71240	71339:	gap of	100	bp	
*	71340	83903:	contly of	12664	bp	in length
*	83904	84003:	gap of	100	bp	
*	84004	91357:	contly of	7354	bp	in length
*	91358	91457:	gap of	100	bp	
*	91458	96455:	contly of	4995	bp	in length
*	96453	96552:	gap of	100	bp	
*	96553	148189:	contly of	51637	bp	in length
*	148190	148289:	gap of	100	bp	
*	148280	159778:	contly of	11489	bp	in length
*	159729	159878:	gap of	100	bp	
*	159879	164891:	contly of	5013	bp	in length

FEATURE
Source

1. 164891

```

/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```
/chrchromosome="1"
```

```
/clone="RP11-444C12"  
/clone_lib="RP11-11.2"
```

1. .34236

```

/NOTE="assembly_
fragment_chain:1
close and .T7

```

```

crome_end:1)
vector_side:left"

```

34337.47372

```
/note="assembly_fragment:00312
fragment_obs:1"
```

11.dagmenc_cha111:1
47473.1.67025

/note="assembly_fragment:01923

```
fragment_chain:1"
```

```
6/126. . /1239
/note="assembly fragment.02094
```

```
fragment_chain:1"
```

71340.83903

```

/note="assembly_fragment:02213
fragment chain:1"

```

84004. .91357

```
/note="assembly_fragment:01292
```

```
fragment_chain:1"
01458 96453
```

/note="assembly_fragment:01391

```
fragment_chain:1"
```

96553. 148189
not a "comb" - fragment - 03015

```

/NOLE= assembly:02043
fragment chain:1"

```

148290.159778

```
/note="assembly_fragment:00427
```

```
11 argument_chain: 1
159879: 164891
```


/note="assembly_fragment:00773
fragment_chain:1
clone_end:SP6
vector_side:right"
BASE COUNT 48381 a 32179 c 32357 g 51067 t 907 others
ORIGIN

Query Match 52.4% Score 143; DB 2; Length 164891;
Best Local Similarity 99.38; Pred. No. 3,3e-29;
Matches 143; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 53 AGGACATGTCAGCCAGCAGCATCTGTATCTGTTTGAAGCTACAGAAAGCATTTTATT 112
|||||
Db 75701 AGGACATGTCAGCCAGCAGCATCTGTATCTGTTTGAAGCTACAGAAAGCATTTTATT 75642
Y 113 TCAGAAATGTTGCCATTTTGTATCTCTGAACATGGAAGACAAAGNTGACTATGTGAGAC 172
|||||
Db 75641 TCAGAAATGTTGCCATTTTGTATCTCTGAACATGGAAGACAAAGCTGACTATGTGAGAC 75582
Y 173 CAAACTTGAGAGCTACAAAATG 196
|||||
Db 75581 CAAACTTGAGAGCTACAAAATG 75558

RESULT 12
HSJ651E10 113764 bp DNA PRI 27-MAY-2000
LOCUS Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1,
DEFINITION complete sequence.
ACCESSION AL122002
VERSION AL122002.16 GI:8247274
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Brown, A.
TITLE Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL Direct Submission
C810 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk
On Jun 4, 2000 this sequence version replaced gi:8247030.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
EM: EMBL; SW: SWISSPROT; TR: TREMBL; WP: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/chr1
RP4-651E10 is from the library RP4-4 constructed at the Roswell
Park Cancer Institute by the group of Pletcher de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP4-651E10.

FEATURES
Source
1. 113764
/organism="Homo sapiens"

/db_xref="taxon:9606"
/chromosome="1"
/map="p22.3-31.1"
/clone="RP4-651E10"
/clone_1b="RP4-4"
BASE COUNT 34890 a 21989 c 22351 g 34534 t
ORIGIN

Query Match 51.8% Score 141.4; DB 9; Length 113764;
Best Local Similarity 98.68; Pred. No. 8.8e-29;
Matches 142; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Y 53 AGGACATGTCAGCCAGCAGCATCTGTATCTGTTTGAAGCTACAGAAAGCATTTTATT 112
|||||
Db 6385 AGGACATGTCAGCCAGCAGCATCTGTATCTGTTTGAAGCTACAGAAAGCATTTTATT 6444
Y 113 TCAGAAATGTTGCCATTTTGTATCTCTGAACATGGAAGACAAAGNTGACTATGTGAGAC 172
|||||
Db 6445 TCAGAAATGTTGCCATTTTGTATCTCTGAACATGGAAGACAAAGCTGACTATGTGAGAC 6504
Y 173 CAAACTTGAGAGCTACAAAATG 196
|||||
Db 6505 CAAACTTGAGAGCTACAAAATG 6528

RESULT 13
AF001264 1265 bp mRNA MAM 18-NOV-1997
LOCUS Bos taurus clone 4 endothelial adhesion molecule Lu-BCAM-1 mRNA,
DEFINITION complete cds.
ACCESSION AF001264
VERSION AF001264.1 GI:2623768
KEYWORDS
SOURCE COW.
ORGANISM Bos taurus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
AUTHORS Elble, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
Goodwin, A. and Pauli, B.U.
TITLE Cloning and characterization of Lu-BCAM-1 suggest it is an
JOURNAL endothelial chloride channel
REFERENCE Unpublished
2 (bases 1 to 1265)
AUTHORS Elble, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
Goodwin, A. and Pauli, B.U.
TITLE Direct Submission
JOURNAL Submitted (25-APR-1997) Vet. Pathology, Cornell U., VMC C4-146,
Ithaca, NY 14853, USA

FEATURES
Source
1. 1265
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="lung"
/cell_type="aortic endothelial cells"
/clone="4"
109. 1137
/note="endothelial adhesion molecule; chloride channel;
secreted form"
/codon_start=1
/product="Lu-BCAM-1"
/protein_id="AAB6532.1"
/db_xref="GI:2623768"
/translation="MWLCNVILFLITLHLLPKKSSVNLINNGYDGIIVAINPSVE
DEKILEIKEMVEASTYTLPHARRRYERNVSLIPMTWKSSEYFIPQESYDAD
VIVANPIKYGDDPYTLQYRCGRGKRYIHFTNPLLTNNPHYSGRQVFEWMAH
RMGIJEDYNDQDPYISRKNTLEATRCSTHITGIVNFKKPGGSCITSLCRDSQSG
LYKACFEFLPKSOTAKESIMPELSHSVTEFEKTHNTEAPLQKMGKSTWV
INMSVDQNTSPMTENMPPIHPFISLKSQRVYCVLDKSGMSADITLLALINI
FKLIGNIT"

BASE COUNT 440 a 231 c 231 g 363 t

ORIGIN

Query Match 45.6%; Score 124.6; DB 4; Length 1265;
Best Local Similarity 65.7%; Pred. No. 2.5e-24;
Matches 178; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

0y 2 TTGCAATGACCCCAATGTGCGAAGATGAACACTATTCACAAATTAAGACATG 61
Db 215 TTGCAATTAACCCAGTGTGCGAAGATGAACAACTATTCACAAATTAAGACATG 274
0y 62 TGACCCAGGACCTCTGTATCTGTGTTGAAGCTACAGGAAAGGATTTTTCAAAATG 121
Db 275 TTAACGAAAGCTTCTACTTACTTACTTCTGATCCACCAAGAGAGTTTATTCAGAGATG 334
0y 122 TTGCAATTTGATTCCTGTAAGAGCAAGAGNTGATGATGACCAAAATCTG 181
Db 335 TGAGCAATTTTATTCATTAATGACCTGGAATCAAAATCTGAGTCTTCAATCAAAATG 394
0y 182 AGACCTACAAAATGCTGATGTTGCTGATGCTGATCTGCTCCAGGAAATGATGAC 241
Db 395 AATCATATGACCAAGCAGATGTCATGTTGCTATTCCTATTAATAATGAGATGATC 454
0y 242 CCTACACTGNGCAGATGGCGAAGCTGTGCGA 272
Db 455 CCTATACACTTCAATATGGAAGGTGTGAGA 485

RESULT 14

AF001263 2820 bp mRNA MAM 18-NOV-1997
LOCUS Bos taurus clone 3 endothelial adhesion molecule Lu-ECAM-1 mRNA,
DEFINITION complete cds.
ACCESSION AF001263
VERSION AF001263.1 GI:2623766

SOURCE

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 2820)
Eblie, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
Goodwin, A., and Pauli, B.U.
Cloning and characterization of Lu-ECAM-1 suggest it is an
endothelial chloride channel

TITLE

Unpublished
2 (bases 1 to 2820)
Eblie, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
Goodwin, A., and Pauli, B.U.
Direct Submission
Submitted (25-APR-1997) Vet. Pathology, Cornell U., VMC C4-146,
Ithaca, NY 14853, USA

FEATURES

source
1..2820
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="3"
/tissue_type="lung"
/cell_type="aortic endothelial cells"
194..2656
/note="endothelial adhesion molecule; chloride channel"
/codon_start=1
/product="Lu-ECAM-1"
/db_xref="GI:2623767"

CDS

1..2820
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="3"
/tissue_type="lung"
/cell_type="aortic endothelial cells"
194..2656
/note="endothelial adhesion molecule; chloride channel"
/codon_start=1
/product="Lu-ECAM-1"
/db_xref="GI:2623767"

KAGFOAIHSDSTSGSEIILLTDEEDNEINSCFEDVKRSAGIHTTALGPSANELE
TLSNMTGVRFPFANDITGLTNAFERSISRSISQIOALQLESKALKITGRKRVNGTV
PVDSTVGDTEFVWTLQKPEIVLQDPKGRKKTSDPEEDLUNRSARLOJPGIAET
GFTVYSLINLNASOMLTVTWTTRARSPITPVIATAHASHOHTAAYSPMIVYAGVSO
GFLPVLGSLVIAIIETEDGHOVTELEMDNGARDPYKNDGITSRTFTDYGGRSLK
VHAQRNNATARNLQKPNKLYVGVYENGTILNPPREVKDOLAKAKIEDFSRLT
SGSFETVAGAPPPGNPSVFPSPKLTIDLEAKREKDYIOLSWTAPQNVLDKRGASFP
MSRFSHQVAKYLEIOLQHSFO"

Query Match 45.6%; Score 124.6; DB 4; Length 2820;
Best Local Similarity 65.7%; Pred. No. 2.8e-24;
Matches 178; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

0y 122 TTGCAATGACCCCAATGTGCGAAGATGAACACTATTCACAAATTAAGACATG 61
Db 300 TTGCAATTAACCCAGTGTGCGAAGATGAACAACTATTCACAAATTAAGACATG 359
0y 62 TGACCCAGGACCTCTGTATCTGTGTTGAAGCTACAGGAAAGGATTTTTCAAAATG 121
Db 360 TTAACGAAAGCTTCTACTTACTTACTTCTGATCCACCAAGAGAGTTTATTCAGAGATG 419
0y 182 AGACCTACAAAATGCTGATGTTGCTGATGCTGATCTGCTCCAGGAAATGATGAC 241
Db 420 TGAGCAATTTTATTCATTAATGACCTGGAATCAAAATCTGAGTCTTCAATCAAAATG 479
0y 242 CCTACACTGNGCAGATGGCGAAGCTGTGCGA 272
Db 540 CCTATACACTTCAATATGGAAGGTGTGAGA 570

RESULT 15

AF001262 3288 bp mRNA MAM 18-NOV-1997
LOCUS Bos taurus clone 2 endothelial adhesion molecule Lu-ECAM-1 mRNA,
DEFINITION complete cds.
ACCESSION AF001262
VERSION AF001262.1 GI:2623764

SOURCE

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 3288)
Eblie, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
Goodwin, A., and Pauli, B.U.
Cloning and characterization of Lu-ECAM-1 suggest it is an
endothelial chloride channel

TITLE

Unpublished
2 (bases 1 to 3288)
Eblie, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
Goodwin, A., and Pauli, B.U.
Direct Submission
Submitted (25-APR-1997) Vet. Pathology, Cornell U., VMC C4-146,
Ithaca, NY 14853, USA

FEATURES

source
1..3288
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="2"
/tissue_type="lung"
/cell_type="aortic endothelial cells"
59..2443
/note="endothelial adhesion molecule; chloride channel"
/codon_start=1
/product="Lu-ECAM-1"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

```
Run on: April 4, 2002, 03:17:07 ; Search time 1321.64 Seconds
        (without alignments)
        177.090 Million cell updates/sec
```

Title: US-09-049-696-2
 Perfect score: 273
 Sequence: 1 GTTGCATTGACCCCAATGTAGATTGGCAACTGTGGCAG 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      930621 segs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
_N_Geneseq_1101.*
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	267.4	97.9	2742	22	AAH46102	Human CLC1A1 coding
2	267.4	97.9	2745	20	AAAF81927	Human ICACC-1 nucl
3	267.4	97.9	2825	22	AAH46124	Human CLC1A1 cDNA,
4	267.4	97.9	2854	22	AAH381879	Human colon cancer
5	267.4	97.9	2854	22	AAAF81787	Human colon cancer
6	267.4	97.9	2867	22	AAH33285	Human secreted pro
7	267.4	97.9	3111	20	AAZ09840	Human colon cancer
8	267.4	97.9	3311	22	AAI29502	Human membrane spa
9	245.4	89.9	401	21	AAAF77755	C902P determined c
10	245.4	89.9	401	21	AAI28493	cDNA encoding huma
11	210.4	77.1	3109	22	AAH35019	Colon tumour relat
						Human colon cancer

1	12	183.2	66.7	2739	22	AAH46101	Mouse Gob-5 coding
2	13	183.2	66.7	2843	22	AAH46120	Mouse Gob-5 cDNA.
3	14	182.2	66.7	2931	20	AAE81925	Murine ICAC-1 nuc
4	15	151.2	55.4	3265	21	AAZ65095	Membrane-bound pro
5	16	151.2	55.4	3265	21	AAE92092	Human PRO1124 cDNA
6	17	151.2	55.4	3265	22	AAE44241	Human PRO1124 (UNQ
7	18	108	39.6	375	22	AAE65887	Novel human polyonu
8	19	75.2	27.5	2784	20	AAZ24658	Human lung tumor a
9	20	75.2	27.5	2784	21	AAE65897	Human lung cancer-
10	21	75.2	27.5	3156	21	AAE51625	Human membrane cha
11	22	75.2	27.5	3362	20	AAZ24657	Human lung tumor a
12	23	75.2	27.5	3362	21	AAE65896	Human lung cancer-
13	24	75.2	27.5	3951	20	AAZ24653	Human lung tumor a
14	25	75.2	27.5	3951	21	AAE65892	Human lung cancer-
15	26	75.2	27.5	8031	21	AAE65950	Human lung cancer-
16	27	71.4	26.2	3190	20	AAE81926	Human lung cancer-
17	28	54.8	20.1	352	21	AAE03649	Human ICAC-2 nucl
18	29	54.8	20.1	398	20	AAH40511	Human secreted pro
19	30	48.6	17.8	227	22	AAH70872	Human cervical can
20	31	48.6	17.8	350	22	AAH72429	Human cervical can
21	32	48.6	17.8	350	22	AAH72994	Human cervical can
22	33	48.6	17.8	454	22	AAH69537	Human cervical can
23	34	34.6	12.7	3044	22	AAH54926	S. epidermidis gen
24	35	34.6	12.6	1200	14	AAO52267	Sequence encoding
25	36	33.6	12.3	3146	22	AAH54011	S. epidermidis gen
26	37	33.6	12.3	3569	22	AAH54940	S. epidermidis gen
27	38	33.2	12.2	799	20	AAH13748	Enterococcus faeca
28	39	33.2	12.2	936	22	AAE58252	Oligonucleotide D1
29	40	33.2	12.2	936	22	AAE58254	Oligonucleotide D1
30	41	33.2	12.2	936	22	AAE58257	Oligonucleotide D1
31	42	33.2	12.2	936	22	AAE58259	Oligonucleotide D1
32	43	33.2	12.2	936	22	AAE58259	Oligonucleotide D2
33	44	33.2	12.2	938	22	AAE58255	Oligonucleotide D1
34	45	33.2	12.2	2557	21	AAAC42981	Arabidopsis thalia

ALIGNMENTS

RESULT 1

1D AAH46102 standard; DNA; 2742 BP

AC AAH46102

11-SEP-2001 (first entry)

Human CLCA1 coding sequence, SEQ ID NO:4.

KW Huijien CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;

chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

XX
XX
XX

Est	CLS	Est	CLS
1.2742	1.2742	1.2742	1.2742

33

/note= "No stop codon given in the specification"

PN WC:200138530-AL
YX

31, MAY - 2001.

22 NOV - 2000; 2000MO-JF006232Z
 PP
 XX

27-APR-2000: 2000JP-0127589

DAI TAKEDA CHEM IND LTD

XX Nakanishi A, Morita S;
PI WPI; 2001-355935/37.
DR P-PSDB; AAB/3716.
XX
PT New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease
PS Claim 3; Page 82-84; 104pp; Japanese.
XX
CC The invention relates to an antisense nucleotide targeted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the Gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
CC pulmonary disease. The present sequence represents the human CLCA1
CC gene coding sequence.
XX
SQ Sequence 2742 BP; 833 A; 616 C; 623 G; 670 T; 0 other:
Query Match 97.9%; Score 267.4; DB 22; Length 2742;
Best Local Similarity 98.2%; Pred. No. 2.4e-72;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GTTGCAATCGACCCCAATGTGCGAGAGATGAAACATCTATTCACAAATTAAGGACATG 60
DB 109 gtgcaatcgaccccaatgtgcgagaagatgaacacatcacaataaagacatg 168
QY 61 GTGACCCAGGATCTCTGATCTGTTGAACCTACAGCAAAAGCATTTATTCAAAAT 120
DB 169 gtgacccaggaatctctgatactgttgaagctacagaagcgaatttatttcaaaaat 228
QY 121 GTTGCAATTTGATTTCCGAAACATGGAAGCAAGTACTATGTGAGACCAAACTT 180
DB 229 gtgcaatcttgatctcctgaacatggaagcaagcgtgactatgtgagacaaaactt 288
QY 181 GAGACCTACAAAATGCTGATGTTCTGCTGAGTCTGAGTCTTCTCCAGGNAATGATGAA 240
DB 289 gagacctacaataatgctgatactgttctgtgctactcctccaggaatgatgaa 348
QY 241 CCTTCACTGNGCAGATGGGCAACTGTGGCAG 273
DB 349 ccttcaactgagcagatgagcaactgtgagag 381
RESULT 2
AAH81927
ID AAH81927 standard; cDNA; 2745 BP.
XX
AC AAH81927;
XX
DT 13-JUN-2001 (first entry)
XX
DE Human ICACC-1 nucleotide sequence.
XX
KW ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
KW interleukin 9 induced calcium activated chloride channel; IL-9;
KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
KW anti-inflammatory; immunomodulatory; cystic fibrosis;
KW inflammatory bowel disease; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2745
FT /*tag= a
FT /product= "ICACC-1"

FT /note= "IL-9 induced calcium activated chloride channel"
XX
XX MO3944620-A1.
PN
XX
XX 10-SEP-1999.
PD
XX
XX 03-MAR-1999; 99WO-US04703.
PF
XX
XX 03-MAR-1998; 98US-0076815.
PR
XX
XX (MGA-) MAGALIN PHARM INC.
PA
XX
XX Holroyd KJ, Levitt RC, Maloy WL, Lounshed J, McLane M;
PI Nicotinic acid, Zhou Y, Dong O;
PI
XX
XX WPI; 1999-550979/46.
DR
DR P-PSDB; AAB/4824.
XX
XX
XX New nucleic acid encoding calcium activated chloride channel, used to
PT identify, e.g. specific modulators for treating atopic allergy -
PT
XX
XX Claim 1; Fig 4B; 75pp; English.
PS
XX
XX The present sequence encodes the human interleukin 9 (IL-9) induced
CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
CC have anti-allergic, anti-asthmatic, anti-inflammatory and
CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
CC used to alleviate asthma (or more generally atopic allergy), while those
CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma
CC (increased levels) or IBD (reduced levels), also for monitoring
CC treatment of these conditions. The ICACC proteins can be used:
CC (i) to raise specific antibodies (Ab); (ii) as fragments or mutants, as
CC reagents, and (iii) as therapeutic (A); (iv) as fragments or mutants, as
CC (A) to disrupt bonding between ICACC and its ligands; and (iii) to
CC identify modulators and binding partners. ICACC polynucleotides can be
CC used to generate transgenic animals or recombinant cells, used to screen
CC for antagonists, also as a source of therapeutic antisense agents or
CC diagnostic probes (for quantifying mRNA expression, e.g. for
CC identification of modulators).
XX
SQ Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other:
Query Match 97.9%; Score 267.4; DB 20; Length 2745;
Best Local Similarity 98.2%; Pred. No. 2.4e-72;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GTTGCAATCGACCCCAATGTGCGAGAGATGAAACATCTATTCACAAATTAAGGACATG 60
DB 109 gtgcaatcgaccccaatgtgcgagaagatgaacacatcacaataaagacatg 168
QY 61 GTGACCCAGGATCTCTGATCTGTTGAACCTACAGCAAAAGCATTTATTCAAAAT 120
DB 169 gtgacccaggaatctctgatactgttgaagctacagaagcgaatttatttcaaaaat 228
QY 121 GTTGCAATTTGATTTCCGAAACATGGAAGCAAGTACTATGTGAGACCAAACTT 180
DB 229 gtgcaatcttgatctcctgaacatggaagcaagcgtgactatgtgagacaaaactt 288
QY 181 GAGACCTACAAAATGCTGATGTTCTGCTGAGTCTGAGTCTTCTCCAGGNAATGATGAA 240
DB 289 gagacctacaataatgctgatactgttctgtgctactcctccaggaatgatgaa 348
QY 241 CCTTCACTGNGCAGATGGGCAACTGTGGCAG 273
DB 349 ccttcaactgagcagatgagcaactgtgagag 381
RESULT 3
AAH46124
ID AAH46124 standard; cDNA; 2825 BP.

XX	AH46124:
AC	
XX	
DT	11-SEP-2001 (first entry)
XX	
DE	Human CLCA1 cDNA, SEQ ID NO:26.
XX	
KW	Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
KW	expression inhibition; antisense therapy; gene therapy;
KW	chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
RN	ss.
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS 23..2767
FT	/tag= a
FT	/product= "Human CLCA1"
FT	/transl_except= (pos:476..478, aa:Lys)
XX	
PN	WO200138530-A1.
PD	
PD	31-MAY-2001.
PE	
PE	22-NOV-2000; 2000WO-JP08232.
XX	
PR	24-NOV-1999; 99JP-0333479.
PR	27-APR-2000; 2000JP-0127589.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
XX	
PI	Nakanishi A, Morita S;
DR	WPI; 2001-355935/37.
DR	P-PSDB; AAB73716.
XX	
PT	New antisense nucleotide, useful for treatment and prevention of
PT	bronchial asthma and chronic obstructive pulmonary disease
XX	
PS	Example 5; Page 92-94; 104pp; Japanese.
CC	The invention relates to an antisense nucleotide targeted to the mouse
CC	Gob-5 gene (coding sequence shown in AH46101) or its human counterpart,
CC	the CLCA1 gene (coding sequence shown in AH46102). The invention also
CC	relates to an antibody specific for the Gob-5 protein, medical and
CC	diagnostic compositions containing the antisense nucleotide or the
CC	antibody, and methods and kits for screening for compounds which inhibit
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC	The antisense oligonucleotides and antibody are therefore useful for the
CC	treatment and prevention of bronchial asthma and chronic obstructive
CC	pulmonary disease. The present sequence represents human CLCA1 cDNA.
XX	
SX	Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;
Query Match	97.9%; Score 267.4; DB 22; Length 2825;
Best Local Similarity	98.2%; Pred. No. 2.4e-72;
Matches 268; Conservative	0; Mismatches 5; Indels 0; Gaps 0;
Df	
Df	1 GTTGCAATCGAGCCCAATGTCGACAAGATGAAGAATCATCTAACAAATAAGGACATG 60
Df	
Df	131 gtcgaatcgagcccaatgctcagaagaatgaaacattcaacaataaaggacatg 190
Df	
Df	61 GTGACCAGGCATCTCTGTATCTGTTTAAAGCTACAGAAAACGATTATTTCAAAT 120
Df	
Df	191 gtgaccagagcatctcgtatctgttgaagtactaagaaagcgattttattccaatat 250
Df	
Df	121 GTGGCATTGTTGATTCCTGGAACATGGAAGACAAAGNTGACTATGTGAGACCAAACTT 180
Df	
Df	251 gtggcatttgatctctgaaacaatggaagaagaagctgactatgtgagcaaaaactt 310
Df	
Df	181 GAGACCTCAAAAAAATGCTGATGTTGTGGTGTGAGTCTAATTCCTCCGAGNAATATGAA 240
Df	

[illegible]

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
CC
XX
SQ Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;

Query Match 97.9%; Score 267.4; DB 22; Length 2867;
Best Local Similarity 98.2%; Pred. No. 2.4e-72;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 GTTGAATGACCCCAATGTCGAGAGATGAACACTCATTCACAAATTAAGACATG 60
DB 146 gtgcaatgcagcccaatgctgagagatgaacacatcattcaacaataaaggagacatg 205
OY 61 GTGACCCAGGACATCTGTATCTGTTGAAGCTACAGAAAGCGATTTTATTCAAAAT 120
DB 206 gtgaccagagcatctctgtatctgtttgaagctacaggaagagatttatttcaaaaat 265
OY 121 GTTGCCATTTGATTCCTGAACATGGAAGACAAAGGACTATGTGAGACCAAACTT 180
DB 266 gtgacatttgcattctcctgaacaatggaagaagaagctgactatgtgagaccacaactt 325
OY 181 GAGACCTACAAAATGCTGATGTTGCTGAGCTCTANTCCTCCAGGNATGATGAA 240
DB 326 ggaacttacaataaagctgagatgctgtgctgagctactcctcagtaatgaa 385
OY 241 CCTACACTGNCAGATGGCAACTGTGCGAG 273
DB 386 cctacactgagcagatgagcaactgtgagag 418

RESULT 7
AAZ09840
ID AAZ09840 standard; cDNA; 3111 BP.
XX
AC AAZ09840;
XX
DT 26-NOV-1999 (first entry)
XX
DE Human membrane spanning protein MSP-5 cDNA fragment 2.
XX
KW Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;
KM neoplastic disorder; immunological disorder; reproductive disorder;
KM MSP-5; ds.
XX
OS Homo sapiens.
XX
PN WO946380-A2.
XX
PD 16-SEP-1999.
XX
PF 09-MAR-1999; 99WO-US05073.
XX
PR 13-MAR-1998; 98US-003064.
XX

XX (INCYTE) INCYTE PHARM INC.
XX Tang YT, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;
XX Giegler KJ, Kaser MR, Baughn MR, Shah P;
XX WHI; 1999-551409/46.
XX P-PSDB: AAY33298.
XX
XX New human membrane spanning proteins used to, e.g. prevent and treat
XX neoplastic disorders -
XX
XX Example 1; Page 80-81; 81pp; English.
XX
XX This invention describes novel human membrane spanning proteins (MSPs),
XX and the polynucleotides encoding them. The products of the invention are
XX used to diagnose, prevent and treat neoplastic, immunological and
XX reproductive disorders. This sequence encodes a human membrane spanning
XX protein MSP-5 fragment.
XX
XX
SQ Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match 97.9%; Score 267.4; DB 20; Length 3111;
Best Local Similarity 98.2%; Pred. No. 2.5e-72;
Matches 268; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 GTTGAATGACCCCAATGTCGAGAGATGAACACTCATTCACAAATTAAGACATG 60
DB 142 gtgcaatgcagcccaatgctgagagatgaacacatcattcaacaataaaggagacatg 201
OY 61 GTGACCCAGGACATCTGTATCTGTTGAAGCTACAGAAAGCGATTTTATTCAAAAT 120
DB 202 gtgaccagagcatctctgtatctgtttgaagctacaggaagagatttatttcaaaaat 261
OY 121 GTTGCCATTTGATTCCTGAACATGGAAGACAAAGGACTATGTGAGACCAAACTT 180
DB 262 gtgacatttgcattctcctgaacaatggaagaagaagctgactatgtgagaccacaactt 321
OY 181 GAGACCTACAAAATGCTGATGTTGCTGAGCTCTANTCCTCCAGGNATGATGAA 240
DB 322 ggaacttacaataaagctgagatgctgtgctgagctactcctcagtaatgaa 381
OY 241 CCTACACTGNCAGATGGCAACTGTGCGAG 273
DB 382 cctacactgagcagatgagcaactgtgagag 414

RESULT 8
AAI29502
ID AAI29502 standard; cDNA; 3111 BP.
XX
AC AAI29502;
XX
DT 12-OCT-2001 (first entry)
XX
DE C932P determined cDNA sequence.
XX
KW human; immunotherapy; diagnosis; colon cancer; colon tumour;
KM immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200149716-A2.
XX
PD 12-JUL-2001.
XX
PF 29-DEC-2000; 2000WO-US35596.
XX
PR 30-DEC-1999; 99US-0476296.
XX
PR 10-JAN-2000; 2000US-0480321.
XX
PR 15-FEB-2000; 2000US-0504629.
XX
PR 06-MAR-2000; 2000US-0519444.
XX
PR 19-MAY-2000; 2000US-0575251.
XX


```

XX  WO200138530-A1.
XX
XX  31-MAY-2001.
XX
XX  22-NOV-2000: 2000MO-JP08232.
XX
XX  24-NOV-1999: 99QP-0333479.
XX
XX  27-APR-2000: 2000JP-0127589.
XX
XX  (TAKE ) TAKEDA CHEM IND LTD.
XX
XX  Nakanishi A, Morita S;
XX
XX  WPI; 2001-355935/37.
XX
XX  P-PSDB; AAB73715.
XX
XX  New antisense nucleotide, useful for treatment and prevention of
XX  bronchial asthma and chronic obstructive pulmonary disease -
XX
XX  Example 1; Page 89-91; 104pp; Japanese.
XX
XX  The invention relates to an antisense nucleotide targeted to the mouse
XX  Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX  the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX  relates to an antibody specific for the Gob-5 protein, medical and
XX  diagnostic compositions containing the antisense nucleotide or the
XX  antibody, and methods and kits for screening for compounds which inhibit
XX  the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX  The antisense oligonucleotides and antibody are therefore useful for the
XX  treatment and prevention of bronchial asthma and chronic obstructive
XX  pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
XX
XX  Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other:
XX
XX  Query Match 66.7%; Score 182.2; DB 22; Length 2843;
XX  Best Local Similarity 79.0%; Pred. No. 3.6e-46;
XX  Matches 214; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
XX
XX  2 TTGCATTCGACCCCAATGTCGCGAAGATGAAACATCTCATTCACAAATAAAGACATGC 61
XX  124 TCGCCAGACACGACGAGCGGAGAGATGAAAGCCCATCACAACATAAGACATG 183
XX  62 TGACCCAGGACATCTCTGTAATCTGTTGAAGCTACAGGAAGCATTTTATTGAAAATG 121
XX  184 TGACTCAGGCTCCATCCTGCTGTTGAGCTCAGGAGAAAGATTCTACTCAAAATG 243
XX  122 TTGCATTTGATTCCTGAAACATGGAAGACAAAGGNTGACTATGTGAGACCAAACTTG 181
XX  244 TTGCATTTGATTCCTGAAACATGGAAGACAAAGGNTGACTATGTGAGACCAAACTTG 303
XX  182 AGACCTACAAAATGCTGATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
XX  304 AAACCTCAAAAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
XX  242 CCTACACTGNGCAGATGGCACTGTGGCA 272
XX  364 CCTACACGACATATAGAGCATGTGGAGA 394
XX
XX  Db
XX
XX  RESULT 14
XX  AAF81925
XX  ID AAF81925 standard; cDNA; 2931 BP.
XX  XX
XX  AAF81925;
XX
XX  13-JUN-2001 (first entry)
XX
XX  Murine ICACC-1 nucleotide sequence.
XX
XX  ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
XX  Interleukin 9 induced calcium activated chloride channel; IL-9;

```

```

XX  calcium activated chloride channel; anti-allergic; anti-asthmatic;
XX  anti-inflammatory; immunomodulatory; cystic fibrosis;
XX  inflammatory bowel disease; autoimmune disease; ss.
XX
XX  Mus sp.
XX
XX  Location/Qualifiers
XX  Key 8..2749
XX  CD8 /tag- a
XX  FT /product- "ICACC-1"
XX  FT /note- "IL-9 induced calcium activated chloride channel"
XX
XX  WQ9944620-A1;
XX
XX  10-SEP-1999.
XX
XX  03-MAR-1999; 99MO-US04703.
XX
XX  03-MAR-1998; 98US-0076815.
XX
XX  (MAG- ) MAGNIN PHARM INC.
XX
XX  Holroyd KJ, Levitt RC, Maloy WL, Louahed J, McLane M;
XX  Nizolides NC, Zhou Y, Dong Q;
XX  WPI; 1999-550979/46.
XX  P-PSDB; AAB74822.
XX
XX  New nucleic acid encoding calcium activated chloride channel, used to
XX  identify, e.g. specific modulators for treating atopic allergy -
XX  Claim 2; Fig 2; 75pp; English.
XX
XX  The present sequence encodes the murine interleukin 9 (IL-9) induced
XX  calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
XX  have anti-allergic, anti-asthmatic, anti-inflammatory and
XX  immunomodulatory activities. Compounds (A) that downregulate ICACC are
XX  used to alleviate asthma (or more generally atopic allergy), while those
XX  (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
XX  inflammatory bowel disease (IBD) (or other autoimmune diseases).
XX  Mediating levels of ICACC-1 can be used in the diagnosis of asthma
XX  (increased levels) or IBD (reduced levels), also for monitoring
XX  treatment of these conditions. The ICACC proteins can be used:
XX  (1) to raise specific antibodies (Ab), useful: (a) as immunoassay
XX  reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
XX  (A) to disrupt bonding between ICACC and its ligands; and (iii) to
XX  identify modulators and binding partners. ICACC polynucleotides can be
XX  used to generate transgenic animals or recombinant cells, used to screen
XX  for antagonists, also as a source of therapeutic antisense agents or
XX  diagnostic probes (for quantifying mRNA expression, e.g. for
XX  identification of modulators).
XX
XX  Sequence 2931 BP; 867 A; 719 C; 684 G; 661 T; 0 other:
XX
XX  Query Match 66.7%; Score 182.2; DB 20; Length 2931;
XX  Best Local Similarity 79.0%; Pred. No. 3.7e-46;
XX  Matches 214; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
XX
XX  2 TTGCATTCGACCCCAATGTCGCGAAGATGAAACATCTCATTCACAAATAAAGACATGC 61
XX  117 TCGCCAGACACGACGAGCGGAGAGATGAAAGCCCATCACAACATAAGACATG 176
XX  62 TGACCCAGGACATCTCTGTAATCTGTTGAAGCTACAGGAAGCATTTTATTGAAAATG 121
XX  177 TGACTCAGGCTCCATCCTGCTGTTGAGCTCAGGAGAAAGATTCTACTCAAAATG 236
XX  122 TTGCATTTGATTCCTGAAACATGGAAGACAAAGGNTGACTATGTGAGACCAAACTTG 181
XX  244 TTGCATTTGATTCCTGAAACATGGAAGACAAAGGNTGACTATGTGAGACCAAACTTG 296
XX  182 AGACCTACAAAATGCTGATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
XX  304 AAACCTCAAAAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
XX
XX  Db
XX
XX  RESULT 14
XX  AAF81925
XX  ID AAF81925 standard; cDNA; 2931 BP.
XX  XX
XX  AAF81925;
XX
XX  13-JUN-2001 (first entry)
XX
XX  Murine ICACC-1 nucleotide sequence.
XX
XX  ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
XX  Interleukin 9 induced calcium activated chloride channel; IL-9;

```


PR 18-AUG-1998; 98US-0096960.
 PR 18-AUG-1998; 98US-0097022.
 PR 19-AUG-1998; 98US-0097141.
 PR 20-AUG-1998; 98US-0097218.
 PR 24-AUG-1998; 98US-0097661.
 PR 26-AUG-1998; 98US-0097951.
 PR 26-AUG-1998; 98US-0097952.
 PR 26-AUG-1998; 98US-0097954.
 PR 26-AUG-1998; 98US-0097955.
 PR 26-AUG-1998; 98US-0097971.
 PR 26-AUG-1998; 98US-0097974.
 PR 26-AUG-1998; 98US-0097978.
 PR 26-AUG-1998; 98US-0097979.
 PR 26-AUG-1998; 98US-0097986.
 PR 26-AUG-1998; 98US-0098014.
 PR 31-AUG-1998; 98US-0098525.
 PR 16-SEP-1998; 98US-0100634.
 PR 12-JAN-1999; 99US-0115565.
 XX
 PA (GETH) GENENTECH INC.

XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK.
 PI Wood WI, Yuan J;
 XX WPI: 2000-072883/06.
 DR P-PSDB; AAY66749.
 XX

Membrane-bound proteins and related nucleotide sequences -

PS Claim 2; Fig 273; 822pp; English.

XX The invention provides membrane-bound PRO polypeptides and
 CC polynucleotides encoding them. The PRO sequences of the invention were
 CC identified based on extracellular domain homology screening. The PRO
 CC sequences have homology with proteins including LDL receptors, TIE
 CC ligands and various enzymes. The membrane-bound proteins and receptor
 CC molecules are useful as pharmaceutical and diagnostic agents. Receptor
 CC immunoadhesins, for instance, can be used as therapeutic agents to block
 CC receptor-ligand interactions. The membrane-bound proteins can also be
 CC employed for screening of potential peptide or small molecule inhibitors
 CC of the relevant receptor/ligand interaction. The PRO encoding sequences
 CC are useful as hybridization probes, in chromosome and gene mapping and in
 CC the generation of antisense RNA and DNA. PRO nucleic acid sequences
 CC will also be useful for the preparation of PRO polypeptides, especially
 CC by recombinant techniques.

XX Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;

Query Match 55.4%; Score 151.2; DB 21; Length 3265;
 Best Local Similarity 71.7%; Pred. No. 1.3e-36;
 Matches 195; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

OY 2 TTGCAATGACCCCAATGCGCAAGATGCAACACTTCAACAATAAAGGACATCG 61
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 DB 131 ttgtatagatctctagtgccagaagatgaataatgaacaaatagagatgag 190
 OY 62 TGACCCAGGACATCTGTATCTGTTGAGCTACAGAAAGGATTTATTTCAAAATG 121
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 DB 191 tgactacagcttctacgtacctgttgaaagccacagaanaagatttttccaanaatg 250
 OY 122 TTGCCATTTTGTCTCCTGAACATGAGACAAAGGNTGATGTGAGACCACAAACTTG 181
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 DB 251 tatctatattatctcgtgaatggaagaatctcagtaacaaagccaaacatg 310
 OY 182 AGACTTACAAAATGCTGATGTTGCTGCTGAGCTGANTCTCCAGGNAATGAGAGAC 241
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 DB 311 aaacacataaactgcgtatgctatagttgcaccactacactccacaggtagatgaac 370
 OY 242 CTTACACTGNGCAGATGGCAACTGTGGCGAG 273
 DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 DB 371 catacacaaagcagtcacagaatgtgagag 402

Search Completed: April 4, 2002, 03:17:13
 Job time: 60185 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:54:36 ; Search time 521.98 Seconds
(without alignments)
66.818 Million cell updates/sec

Title: US-09-049-696-15

Perfect score: 154
Sequence: 1 ACAGATCTTTCATTCCTAT.....TCCCTTTCCTAATTCATA 154

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	92.9	878	1	US-08-469-667-8
2	143	92.9	878	1	US-08-469-667-8
3	56.6	36.8	618	4	US-09-385-982-24
4	29.4	19.1	678	4	US-09-328-111-26
5	29.4	19.1	678	4	US-09-328-111-26
6	28.6	18.6	845	4	US-09-105-390-14
7	28.6	18.6	845	4	US-09-105-390-14
8	27.6	17.9	1552	2	US-08-752-760A-3
9	27.6	17.9	1552	2	US-08-752-760A-3
10	27.6	17.9	1552	2	US-08-752-760A-3
11	27.6	17.9	1552	2	US-08-752-760A-3
12	27.6	17.9	1552	2	US-08-752-760A-3
13	27.6	17.9	1552	2	US-08-752-760A-3
14	27.6	17.9	1552	2	US-08-752-760A-3
15	27.6	17.9	1552	2	US-08-752-760A-3
16	27.6	17.9	1552	2	US-08-752-760A-3
17	27.6	17.9	1552	2	US-08-752-760A-3
18	27.6	17.9	1552	2	US-08-752-760A-3
19	27.6	17.9	1552	2	US-08-752-760A-3
20	27.6	17.9	1552	2	US-08-752-760A-3
21	27.6	17.9	1552	2	US-08-752-760A-3
22	27.6	17.9	1552	2	US-08-752-760A-3
23	27.6	17.9	1552	2	US-08-752-760A-3
24	27.6	17.9	1552	2	US-08-752-760A-3
25	27.6	17.9	1552	2	US-08-752-760A-3
26	27.6	17.9	1552	2	US-08-752-760A-3
27	27.6	17.9	1552	2	US-08-752-760A-3

28	26.2	17.0	430	4	US-08-905-223-235	Sequence 235, App
29	25.8	16.8	878	1	US-08-328-254-5	Sequence 5, Appl
30	25.8	16.8	10136	1	US-08-353-700-2	Sequence 2, Appl
31	25.8	16.8	10136	5	US-08-353-700-2	Sequence 2, Appl
32	25.8	16.8	24183	4	US-08-943-731-3	Sequence 3, Appl
33	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
34	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
35	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
36	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
37	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
38	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
39	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
40	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
41	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
42	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
43	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
44	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl
45	25.8	16.8	280	6	US-08-943-731-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-08-459-667-8

Sequence 8, Application US/08469667

Patent No. 5733748

GENERAL INFORMATION:

APPLICANT: YU, Guo-Liang
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bairo, Gillfillan, Cecchi,
STREET: Stewart & Olstein
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,667
FILING DATE: 06-JUN-1995
CLASSIFICATION: 336

ATTORNEY/AGENT INFORMATION:

NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-435
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
FAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
US-08-459-667-8

Query Match Score 143; DB 1; Length 878;
Best local Similarity 99.4%; Pred. No. 1.6e-39;
Matches 154; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 ACAGATCTTTTCATTGCTATTCAGGCTGTGATAGTGCATCGAATCAGAAATATCC 60
 |||
 DB 545 ACAGATCTTTTCATTGCTATTCAGGCTGTGATAGTGCATCGAATCAGAAATATCC 604
 |||
 OY 61 AACATTGCAGAGATATCTTTGTTATTCCTCCACAGACTCCGACAGACACCTAGTCTT 120
 |||
 DB 605 AACATTGCAGAGATATCTTTGTTATTCCTCCACAGACTCCGACAGACACCTAGTCTT 664
 |||
 OY 121 GATGAAGCTGCTGCTCTGT-CCTAATATTCATA 154
 |||
 DB 665 GATGAAGCTGCTGCTCTGTGCTTAATATTCATA 699
 |||

RESULT 2

PCT-US95-07289-8
 ; Sequence 8, Application PC/TUS9507289
 ; GENERAL INFORMATION:
 ; APPLICANT: Yu, Guo-Liang
 ; APPLICANT: Rosen, Craig
 ; TITLE OF INVENTION: Colon Specific Genes and Proteins
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
 ; ADDRESSEE: Stewart & Olstein
 ; STREET: 6 Becker Farm Road
 ; CITY: Roseland
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07068-1739
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/07289
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ferraro, Gregory D.
 ; REGISTRATION NUMBER: 36,134
 ; REFERENCE/DOCKET NUMBER: 325800-265
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-994-1700
 ; TELEFAX: 201-994-1744
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 878 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 2..685
 ; PCT-US95-07289-8

Query Match 92.9%; Score 143; DB 5; Length 878;
 Best Local Similarity 99.4%; Pred. No. 1.6e-39;
 Matches 154; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 ACAGATCTTTTCATTGCTATTCAGGCTGTGATAGTGCATCGAATCAGAAATATCC 60
 |||
 DB 545 ACAGATCTTTTCATTGCTATTCAGGCTGTGATAGTGCATCGAATCAGAAATATCC 604
 |||
 OY 61 AACATTGCAGAGATATCTTTGTTATTCCTCCACAGACTCCGACAGACACCTAGTCTT 120
 |||
 DB 605 AACATTGCAGAGATATCTTTGTTATTCCTCCACAGACTCCGACAGACACCTAGTCTT 664
 |||
 OY 121 GATGAAGCTGCTGCTCTGT-CCTAATATTCATA 154
 |||
 DB 665 GATGAAGCTGCTGCTCTGTGCTTAATATTCATA 699
 |||

RESULT 3
 US-09-335-982-24/C
 ; Sequence 24, Application US/09385982
 ; Patent No. 6262334
 ; GENERAL INFORMATION:
 ; APPLICANT: ENDEGE, WILSON O., ET AL.
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; FILE REFERENCE: CCDNA-260XX
 ; CURRENT APPLICATION NUMBER: US/09/385,982
 ; CURRENT FILING DATE: 1999-08-30
 ; EARLIER APPLICATION NUMBER: 09/328,111
 ; EARLIER FILING DATE: 1999-06-08
 ; EARLIER APPLICATION NUMBER: 60/117,393
 ; EARLIER FILING DATE: 1999-01-27
 ; EARLIER APPLICATION NUMBER: 60/098,639
 ; EARLIER FILING DATE: 1998-08-31
 ; NUMBER OF SEQ ID NOS: 544
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 24
 ; LENGTH: 618
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(618)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-385-982-24

Query Match 36.8%; Score 56.6; DB 4; Length 618;
 Best Local Similarity 65.6%; Pred. No. 3.2e-10;
 Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;

OY 1 ACAGATCTTTTCATTGCTATTCAGGCTGTGATAGTGCATCGAATCAGAAATATCC 60
 |||
 DB 234 ACCGACATATTTATTCCTTAAGTATAGTAAAGCATTTGACATCAAAAGTATCC 175
 |||
 OY 61 AACATTGCAGAGATATCTTTGTTATTCCTCCACAGACTCCGACAGACACCTAGTCTT 117
 |||
 DB 114 AACATTGCAGAGATATCTTTGTTATTCCTCCACAGAAATCTGATGACATGATCTTCA 115
 |||
 OY 118 CCTGATGAAGCTGCTGCTCTGTGCTTATA 148
 |||
 DB 114 CCTACTCTACTCTACTCTACTCTACTCTGATA 84
 |||

RESULT 4
 US-09-328-111-26
 ; Sequence 26, Application US/09328111
 ; Patent No. 6262333
 ; GENERAL INFORMATION:
 ; APPLICANT: ENDEGE, WILSON O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astle, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Dertl, Adnan
 ; APPLICANT: Ford, Donna M.
 ; APPLICANT: Lewis, Marcia E.
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Schlegel, Robert
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; FILE REFERENCE: CCD-257 (US)
 ; CURRENT APPLICATION NUMBER: US/09/328,111
 ; CURRENT FILING DATE: 1999-06-08
 ; EARLIER APPLICATION NUMBER: US 60/088,801
 ; EARLIER FILING DATE: 1998-06-10

NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 26
LENGTH: 638
TYPE: DNA
ORGANISM: Homo sapiens
US-09-328-111-26

Query Match 19.1%; Score 29.4; DB 4; Length 638;
Best Local Similarity 66.7%; Pred. No. 0.57; Mismatches 21; Indels 0; Gaps 0;
Matches 42; Conservative 0;

QY 62 ACATTGACGAGTATCTTTGTTATTCCTCCACAGACTCGCAGACACCTAGTCTG 121
111 11 111111 111 11 111111 111 1 1111 1111
DB 394 acaagccttagatctactcttcttctgtcagacagactcaccatcgccttctccta 453

QY 122 ATG 124
11
DB 454 gty 456

RESULT 5

US-09-328-111-66/C
Sequence 66, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Margia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 66
LENGTH: 678
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)...(678)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-66

Query Match 19.1%; Score 29.4; DB 4; Length 678;
Best Local Similarity 66.7%; Pred. No. 0.59; Mismatches 21; Indels 0; Gaps 0;
Matches 42; Conservative 0;

QY 62 ACATTGACGAGTATCTTTGTTATTCCTCCACAGACTCGCAGACACCTAGTCTG 121
111 11 111111 111 11 111111 111 1 1111 1111
DB 242 ACAAGCCTTAGTATCTACTTCTTGTCTGACAGACTTCAACCAATGCTCTATTCCTA 183

QY 122 ATG 124
11
DB 182 GTG 180

RESULT 6

US-09-105-390-14
Sequence 14, Application US/09105390
Patent No. 6288303
GENERAL INFORMATION:
APPLICANT: Rodriguez, Raymond
TITLE OF INVENTION: Rice Beta-Glucanase Enzymes
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,390
FILING DATE: Filed herewith
ELASIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/050,675
FILING DATE: 25-JUN-97
ATTORNEY/AGENT INFORMATION:
NAME: Pelthory, Joanne R.
REGISTRATION NUMBER: P42,995
REFERENCE/DOCKET NUMBER: 2000-0455.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
FLEX:

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 845 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-105-390-14

Query Match 18.6%; Score 28.6; DB 4; Length 845;
Best Local Similarity 57.1%; Pred. No. 1.2; Mismatches 39; Indels 0; Gaps 0;
Matches 52; Conservative 0;

QY 3 AGATCTTTGATTCGATTCAGGCTGTGATGAAGTGTGATCAATCAAAATATCCAA 62
1 11111 111 111111 111 111111 111 111111 111111
DB 46 AGCTCTGTGCTATATATGCTGTGCTGCTGCTGTGTTGAACACTAGTGAATCCAG 75
1 11111 111 111111 111 111111 111 111111 111111
QY 53 CATTGACGAGTATCTTTGTTATTCCTCCA 93
1 11111 111 111111 111 111111 111 111111 111111
DB 46 CAAGGACCAAGATCTCAGTTGATGCTTCA 106
1 11111 111 111111 111 111111 111 111111 111111

RESULT 7

US-09-105-390-6
Sequence 6, Application US/09105390
Patent No. 6288303
GENERAL INFORMATION:
APPLICANT: Rodriguez, Raymond
TITLE OF INVENTION: Rice Beta-Glucanase Enzymes
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

3 AGATCTTTTCATTGCTATTTCAGGCTGTTGATAAGGTCGATCTGAATCAGAAATATCCAA 62

QY 127 ACGTCTGCTCCTTGTCT 144
- - - - -
Db. 327 CCTCCTCCTCCTGCTCT 380

RESULT 15

US-08-701-380-1/c

Sequence 1, Application US/08701380

Patent No. 5686598

GENERAL INFORMATION:

APPLICANT: NORTH, Michael

APPLICANT: NISHINA, Patsy

APPLICANT: NAGERT, Juergen

TITLE OF INVENTION: GENES ASSOCIATED WITH RETINAL

TITLE OF INVENTION: DYSTROPHIES

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fleht, Hohbach, Test, Albritton & Herbert

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/701,380

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SHERWOOD, Pamela J.

REGISTRATION NUMBER: 36,677

REFERENCE/DOCKET NUMBER: A-63565/PJS

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-494-8700

TELEFAX: 415-494-8771

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2116 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-701-380-1

Query Match

Best Local Similarity 17.4%; Score 26.8; DB 1; Length 2116;

Matches 46; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

QY 67 GCACGATATCTTGTATTCTCCACACACTCCGCCAGACACCTAGTCCGATGAA 126
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 457 GAAGGATTTCTTCTTCTTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 398
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 127 ACCTCTGCTCTGCTCT 144
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 397 CCTCTCTCTCTCTCTCTCT 380
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  
```

Search completed: April 3, 2002, 20:54:41
Job time: 41839 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:18:21 ; Search time 1321.64 Seconds
(without alignments)
99.897 Million cell updates/sec

Title: US-09-049-696-15

Perfect score: 154
Sequence: 1 ACAGATCTTTTCATTGCTAT.....TCCCTGTCCTAATATTCATA 154

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

N.Geneseq_1101.*
1: /SIDSI/gcgdata/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	154	100.0	2742	22 AAH46102 Human C1CA1 coding
2	154	100.0	2745	20 AAF81927 Human ICACC-1 nucl
3	154	100.0	2825	22 AAH46124 Human C1CA1 CDNA
4	154	100.0	2854	22 AAH34879 Human colon cancer
5	154	100.0	2854	22 AAF81787 Human secreted pro
6	154	100.0	2867	22 AAH33285 Human colon cancer
7	154	100.0	3109	22 AAH35019 Human colon cancer
8	154	100.0	3111	20 AAZ09840 Human membrane spa
9	154	100.0	3311	22 AA129502 C902P determined c
10	143	92.9	878	18 AAT45884 Human colon specif
11	143	92.9	878	19 AAV16672 Polynucleotide seq

12	112	72.7	255	16 AAT22483
13	73.6	47.8	2739	22 AAH46101
14	73.6	47.8	2843	22 AAH46120
15	73.6	47.8	2931	20 AAF81925
16	62.4	40.5	1802	21 AAC98067
17	62.4	40.5	1802	22 AAH3192
18	56.6	36.8	618	21 AAA16019
19	56.6	36.8	2616	21 AAA64335
20	56.6	36.8	3265	22 AAF92092
21	56.6	36.8	3265	22 AAF92092
22	56.6	36.8	3265	22 AAF92092
23	34.4	22.3	2454	22 AAZ24658
24	34.4	22.3	2784	20 AAC24658
25	34.4	22.3	2784	21 AAC65897
26	34.4	22.3	3156	21 AAZ51625
27	34.4	22.3	3362	20 AAZ24657
28	34.4	22.3	3362	20 AAC65896
29	34.4	22.3	3362	20 AAC65896
30	34.4	22.3	3362	20 AAC65896
31	34.4	22.3	3362	20 AAC65896
32	34.4	22.3	3362	20 AAC65896
33	34.4	22.3	3362	20 AAC65896
34	34.4	22.3	3362	20 AAC65896
35	34.4	22.3	3362	20 AAC65896
36	34.4	22.3	3362	20 AAC65896
37	34.4	22.3	3362	20 AAC65896
38	34.4	22.3	3362	20 AAC65896
39	34.4	22.3	3362	20 AAC65896
40	34.4	22.3	3362	20 AAC65896
41	34.4	22.3	3362	20 AAC65896
42	34.4	22.3	3362	20 AAC65896
43	34.4	22.3	3362	20 AAC65896
44	34.4	22.3	3362	20 AAC65896
45	34.4	22.3	3362	20 AAC65896

ALIGNMENTS

1	AAH46102	standard; DNA; 2742 BP.
2	AAH46102	standard; DNA; 2742 BP.
3	AAH46102	standard; DNA; 2742 BP.
4	AAH46102	standard; DNA; 2742 BP.
5	AAH46102	standard; DNA; 2742 BP.
6	AAH46102	standard; DNA; 2742 BP.
7	AAH46102	standard; DNA; 2742 BP.
8	AAH46102	standard; DNA; 2742 BP.
9	AAH46102	standard; DNA; 2742 BP.
10	AAH46102	standard; DNA; 2742 BP.
11	AAH46102	standard; DNA; 2742 BP.
12	AAH46102	standard; DNA; 2742 BP.
13	AAH46102	standard; DNA; 2742 BP.
14	AAH46102	standard; DNA; 2742 BP.
15	AAH46102	standard; DNA; 2742 BP.
16	AAH46102	standard; DNA; 2742 BP.
17	AAH46102	standard; DNA; 2742 BP.
18	AAH46102	standard; DNA; 2742 BP.
19	AAH46102	standard; DNA; 2742 BP.
20	AAH46102	standard; DNA; 2742 BP.
21	AAH46102	standard; DNA; 2742 BP.
22	AAH46102	standard; DNA; 2742 BP.
23	AAH46102	standard; DNA; 2742 BP.
24	AAH46102	standard; DNA; 2742 BP.
25	AAH46102	standard; DNA; 2742 BP.
26	AAH46102	standard; DNA; 2742 BP.
27	AAH46102	standard; DNA; 2742 BP.
28	AAH46102	standard; DNA; 2742 BP.
29	AAH46102	standard; DNA; 2742 BP.
30	AAH46102	standard; DNA; 2742 BP.
31	AAH46102	standard; DNA; 2742 BP.
32	AAH46102	standard; DNA; 2742 BP.
33	AAH46102	standard; DNA; 2742 BP.
34	AAH46102	standard; DNA; 2742 BP.
35	AAH46102	standard; DNA; 2742 BP.
36	AAH46102	standard; DNA; 2742 BP.
37	AAH46102	standard; DNA; 2742 BP.
38	AAH46102	standard; DNA; 2742 BP.
39	AAH46102	standard; DNA; 2742 BP.
40	AAH46102	standard; DNA; 2742 BP.
41	AAH46102	standard; DNA; 2742 BP.
42	AAH46102	standard; DNA; 2742 BP.
43	AAH46102	standard; DNA; 2742 BP.
44	AAH46102	standard; DNA; 2742 BP.
45	AAH46102	standard; DNA; 2742 BP.

XX Nakanishi A, Morita S;
XX WPI: 2001-355935/37.
DR P-PSDB: AAB73716.
XX New antisense nucleotide, useful for treatment and prevention of
PT bronchial asthma and chronic obstructive pulmonary disease
XX
XX Claim 3: Page 82-84; 104pp; Japanese.
XX
XX The invention relates to an antisense nucleotide targeted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the Gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
CC pulmonary disease. The present sequence represents the human CLCA1
CC gene coding sequence.
XX
SQ Sequence 2742 BP: 833 A; 616 C; 623 G; 670 T; 0 other;

Query Match 100.0%; Score 154; DB 22; Length 2742;
Best Local Similarity 100.0%; Pred. No. 1.3e-40;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACAGATCTTTTCATTCGATTCAGGCTGTGTGTAAGTCGATCGAATCAAAATATCC 60
DB 2512 acagatcttttcattgcattcagcgtgtgataagtcgacatcgaataatccc 2571
OY 61 AACATTGCACGAGTATCTTTGTTTTCCTCCACAGACTCCGCCAGAGACACTAGTCT 120
DB 2572 aacattgcacgagatctcttctgttattctccacagactccgccagagacactagctct 2631
OY 121 GATGAACGCTGCTGCTGCTGCTGCTTAATATTCATA 154
DB 2632 gatgaacgctgcctcctctgtcctaataattcata 2665

RESULT 2
AAF81927 ID AAF81927 standard; cDNA: 2745 BP.
XX
AC AAF81927;
XX
DT 13-JUN-2001 (first entry)
XX
XX Human ICACC-1 nucleotide sequence.
XX
KW ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
KW Interleukin 9 induced calcium activated chloride channel; IL-9;
KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
KW anti-inflammatory; immunomodulatory; cystic fibrosis;
KW inflammatory bowel disease; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..2745
FT CDS
FT /*tag= a
FT /product= "ICACC-1"
FT /note= "IL-9 induced calcium activated chloride channel"

XX
XX MO9944620-A1.
XX
XX 10-SEP-1999.
XX
XX 03-MAR-1999; 99WO-US04703.

PR 03-MAR-1998; 98US-0076815.
XX
XX (MAGA-) MAGAIMIN PHARM INC.
XX
XX Heiroyd KJ, Levitt RC, Maloy WL, Louahed J, McLane M;
PI Nigolades NC, Zhou Y, Dong Q;
DR WPI: 1999-550979/46.
DR P-PSDB: AAB74824.
XX
XX New nucleic acid encoding calcium activated chloride channel, used to
PT identify, e.g. specific modulators for treating atopic allergy -
XX
XX Claim 1: Fig 4B: 75pp; English.
XX
XX The present sequence encodes the human interleukin 9 (IL-9) induced
CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
CC have anti-allergic, anti-asthmatic, anti-inflammatory and
CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
CC used to alleviate asthma (or more generally atopic allergy), while those
CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma
CC (increased levels) or IBD (reduced levels), also for monitoring
CC treatment of these conditions. The ICACC proteins can be used:
CC (1) to raise specific antibodies (Ab), useful: (a) as immunoassay
CC reagents, and (b) as therapeutic (A); (11) as fragments or mutants, as
CC (A) to disrupt bonding between ICACC and its ligands; and (11) to
CC identify modulators and binding partners. ICACC polynucleotides can be
CC used to generate transgenic animals or recombinant cells, used to screen
CC for antagonists, also as a source of therapeutic antisense agents or
CC diagnostic probes (for quantifying mRNA expression, e.g. for
SQ identification of modulators).

Sequence 2745 BP: 833 A; 616 C; 625 G; 671 T; 0 other;

Query Match 100.0%; Score 154; DB 20; Length 2745;
Best Local Similarity 100.0%; Pred. No. 1.4e-40;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACAGATCTTTTCATTCGATTCAGGCTGTGTGTAAGTCGATCGAATCAAAATATCC 60
DB 2512 acagatcttttcattgcattcagcgtgtgataagtcgacatcgaataatccc 2571
OY 61 AACATTGCACGAGTATCTTTGTTTTCCTCCACAGACTCCGCCAGAGACACTAGTCT 120
DB 2572 aacattgcacgagatctcttctgttattctccacagactccgccagagacactagctct 2631
OY 121 GATGAACGCTGCTGCTGCTGCTGCTTAATATTCATA 154
DB 2632 gatgaacgctgcctcctctgtcctaataattcata 2665

RESULT 3
AAH46124 ID AAH46124 standard; cDNA: 2825 BP.
XX
AC AAH46124;
XX
DT 11-SEP-2001 (first entry)
XX
XX Human CLCA1 cDNA, SEQ ID NO:26.
XX
KW Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
KW expression inhibition; antisense therapy; gene therapy;
KW chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
KW ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 23..2767
FT CDS


```

FT      /*tag= a
FT      /product= "Human CLCA1"
FT      /transl_except= (pos:476..478, aa:Lys)
XX
XX      WO200138530-A1.
XX
XX      31-MAY-2001.
XX
XX      22-NOV-2000; 2000WO-JP08232.
XX
XX      24-NOV-1999; 99JP-0333479.
XX      27-APR-2000; 2000JP-0127589.
XX
XX      (TAKE ) TAKEDA CHEM IND LTD.
XX
XX      Nakanishi A, Morita S;
XX
XX      WPI: 2001-355935/37.
XX      P-PSDB; AAB73716.
XX
XX      New antisense nucleotide, useful for treatment and prevention of
XX      bronchial asthma and chronic obstructive pulmonary disease -
XX
XX      Example 5; Page 92-94; 104pp; Japanese.
XX
XX      The invention relates to an antisense nucleotide targeted to the mouse
XX      Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX      the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX      relates to an antibody specific for the Gob-5 protein, medical and
XX      diagnostic compositions containing the antisense nucleotide or the
XX      antibody, and methods and kits for screening for compounds which inhibit
XX      the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX      The antisense oligonucleotides and antibody are therefore useful for the
XX      treatment and prevention of bronchial asthma and chronic obstructive
XX      pulmonary disease. The present sequence represents human CLCA1 cDNA.
XX
XX      Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;
XX
XX      Query Match      100.0%; Score 154; DB 22; Length 2825;
XX      Best Local Similarity 100.0%; Pred. No. 1.4e-40;
XX      Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1 ACAGATCTTTGATGCTATTCAGCGCTGTGATAGTCGATCGAATGAAATATCC 60
XX      |||||||
XX      DB      2534 acagatctttcattcgtatcagctgtgataagtcgatcgaatacagaataatcc 2593
XX
XX      QY      61 AACATTGCACGAGATCTTTGTTTATTCCTCCACAGACTCCGCCAGACACCTAGTCCT 120
XX      |||||||
XX      DB      2594 aacattgcacgagatcttctgttattcccccacagactccgcagacactagctct 2653
XX
XX      QY      121 GATGAAGCTGCTGCTCTGCTCTAATATTCATA 154
XX      |||||||
XX      DB      2654 gatgaagctgctgctctgtcctctaataatcata 2687
XX
XX      RESULT 4
XX      AAH34879
XX      ID      AAH34879 standard; cDNA; 2854 BP.
XX
XX      AAH34879;
XX
XX      03-SEP-2001 (first entry)
XX
XX      Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
XX
XX      Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX      colorectal carcinoma; chromosome 1; ss.
XX
XX      Homo sapiens.
XX      WO200122920-A2.
XX

```

```

PD      05-APR-2001.
XX
XX      28-SEP-2000; 2000WO-US26524.
XX
XX      29-SEP-1999; 99US-0157137.
XX      03-NOV-1999; 99US-0163280.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      Rugen SM, Barash SC, Birse CE, Rosen CA;
XX
XX      WPI: 2001-235357/24.
XX      P-PSDB; AAG75474.
XX
XX      Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX      useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX      Claim 1; Page 3462-3463; 9803pp; English.
XX
XX      AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX      cancer-associated nucleic acid molecules (N) and proteins (P), where
XX      the proteins are collectively known as colon cancer antigens. The colon
XX      cancer antigens have cytostatic activity and can be used in gene
XX      therapy and vaccine production. N and P may be used in the prevention,
XX      diagnosis and treatment of diseases associated with inappropriate P
XX      expression. For example, N and P may be used to treat disorders
XX      associated with decreased expression by rectifying mutations or deletions
XX      in a patient's genome that affect the activity of P by expressing
XX      inactive proteins or to supplement the patient's own production of P.
XX      Additionally, N may be used to produce the colon cancer-associated P,
XX      by inserting the nucleic acids into a host cell and culturing the cell
XX      to express the proteins. N and P can be used in the prevention, diagnosis
XX      and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX      and AAB7789 represent sequences used in the exemplification of the
XX      present invention.
XX      N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX      missing at time of publication, meaning no sequences are present for
XX      SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX      Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;
XX
XX      Query Match      100.0%; Score 154; DB 22; Length 2854;
XX      Best Local Similarity 100.0%; Pred. No. 1.4e-40;
XX      Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1 ACAGATCTTTGATGCTATTCAGCGCTGTGATAGTCGATCGAATGAAATATCC 60
XX      |||||||
XX      DB      2546 acagatctttcattcgtatcagctgtgataagtcgatcgaatacagaataatcc 2605
XX
XX      QY      51 AACATTGCACGAGATCTTTGTTTATTCCTCCACAGACTCCGCCAGACACCTAGTCCT 120
XX      |||||||
XX      DB      2606 aacattgcacgagatcttctgttattcccccacagactccgcagacactagctct 2665
XX
XX      QY      121 GATGAAGCTGCTGCTCTGCTCTAATATTCATA 154
XX      |||||||
XX      DB      2656 gatgaagctgctgctctgtcctctaataatcata 2699
XX
XX      RESULT 5
XX      AAF81787
XX      ID      AAF81787 standard; cDNA; 2854 BP.
XX
XX      AAF81787;
XX
XX      12-JUN-2001 (first entry)
XX
XX      Human secreted protein gene 1 SEQ ID NO:11.
XX
XX      Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
XX      dermatological; immunosuppressive; anti-inflammatory; anti-HIV;
XX      immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;
XX      ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
XX

```


KW antialzheimer; antiparkinsonian; antimicrobial; vulnery; gene therapy;
KW immune disorder; hyperproliferative disorder; cardiovascular disease;
KW cancer; angiogenic disorder; neurological disorder; infectious disease;
KW wound healing; regeneration; chemotaxis; chromosome 1; ss.

OS Homo sapiens.

XX WO200112775-A2.

XX 22-FEB-2001.

XX 16-AUG-2000; 2000WO-US22325.

XX 17-AUG-1999; 99US-0149182.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, N1 J, Florence KA, Fiscella M, Wei P, Baker KP;

PI Birse CE, Young PE, Komatsu GA, Moore PA, Soppet DR;

DR WPI; 2001-147550/15.

DR P-PSDB; AAB74733.

XX Nucleic acids encoding 25 human secreted polypeptides, useful for

PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's

PT disease and diabetic retinopathy -

XX Claim 1; Page 441; 485pp; English.

XX AAB74733 to AAB74733 encode the human secreted proteins given in AAB74733

XX to AAB74733. Human secreted proteins can have activities based on the

XX tissues and cells they are expressed in. Example of activities include:

XX immunomodulatory; antisclerotic; dermatological; immunosuppressive;

XX antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;

XX vascular; anti-angiogenic; ophthalmological; neuroprotectant; neurotropic;

XX anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and

XX vulnery. Human secreted proteins can be used in gene therapy and

XX vaccine. Human secreted protein nucleotide sequences (NMI) and proteins

XX (PPI) may be used in the prevention, diagnosis and treatment of diseases

XX associated with inappropriate polypeptide expression. For example, NMI

XX and PPI may be used to treat disorders associated with decreased

XX expression by rectifying mutations or deletions in a patient's genome

XX that affect the activity of proteins by expressing inactive proteins or

XX to supplement the patient's own production of polypeptides. Disorders that

XX may be prevented, diagnosed and/or treated, include immune disorders,

XX hyperproliferative disorders (e.g. cancers), cardiovascular diseases,

XX angiogenic disorders, neurological disorders, infectious diseases and/or

XX for promoting wound healing, regeneration and/or chemotaxis. AAB74733 to

XX AAB74733 and AAB74732 represent sequences used in the exemplification of

XX the present invention.

XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

XX

XX

ID AAB33285 standard; cDNA; 2867 BP.

XX AAB33285;

AC AAB33285;

XX 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO: 341.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX colorectal carcinoma; chromosome 1; ss.

XX Homo sapiens.

XX WO/00122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

PI WPI; 2001-235357/24.

DR P-PSDB; AAG73854.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 1; Page 2452-2453; 9803pp; English.

XX AAB32943 to AAB37195 and AAG73514 to AAG77788 represent human colon

XX cancer-associated nucleic acid molecules (N) and proteins (P), where

XX the proteins are collectively known as colon cancer antigens. The colon

XX cancer antigens have cytostatic activity and can be used in gene

XX therapy and vaccine production. N and P may be used in the prevention,

XX diagnosis and treatment of diseases associated with inappropriate P

XX expression. For example, N and P may be used to treat disorders

XX associated with decreased expression by rectifying mutations or deletions

XX in a patient's genome that affect the activity of P by expressing

XX inactive proteins or to supplement the patient's own production of P.

XX Additionally, N may be used to produce the colon cancer-associated P,

XX by inserting the nucleic acids into a host cell and culturing the cell

XX to express the proteins. N and P can be used in the prevention, diagnosis

XX and treatment of colorectal carcinomas and cancers. AAB37196 to AAB37204

XX and AAB77789 represent sequences used in the exemplification of the

XX present invention.

XX N:3. Pages 666 to 682 and page 7053 of the sequence listing were

XX missing at time of publication, meaning no sequences are present for

XX SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

RESULT 7
ID AAH35019 standard; cDNA; 3109 BP.
XX
AC AAH35019;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:2101.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma; ss.
XX
OS Homo sapiens.
XX
PN M0200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000MO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
XX 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
XX P-PSDB; AAG75614.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 1; Page 3587-3588; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing
XX inactive proteins or to supplement the patient's own production of P.
XX Additionally, N may be used to produce the colon cancer-associated P,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAG77789 represent sequences used in the exemplification of the
XX present invention.
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match 100.0%; Score 154; DB 22; Length 3109;
Best Local Similarity 100.0%; Pred. No. 1.4e-40;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGATCTTTTCATGCTATTCAGAGCTGTGATTAAGTCGATCGAATCAGAAATATCC 60
DB 2399 acagatcttttcattgctattcagcgtgtgataagtcgattcgaatcgaatatcc 2458
QY 61 AACATTGCAGAGATATCTTTGTTTATTCCTCCACAGACTCGCCAGACACCTAGTCT 120
DB 2459 aacattgcagagatattcttatttattcctccacagactcgcagagacacccagtcct 2518
QY 121 GATGAACGTCGTCTCTTCTCTTAATATTCATA 154

bp 2519 gatgaacgctctctcttcttccataatcata 2552
bp 2519 gatgaacgctctctcttcttccataatcata 2552
RESULT 8
ID AA209840 standard; cDNA; 3111 BP.
XX
AC AA209840;
XX
DT 26-NOV-1999 (first entry)
XX
DE Human membrane spanning protein MSP-5 cDNA fragment 2.
XX
KW Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;
KM neoplastic disorder; immunological disorder; reproductive disorder;
XX MSP-5; ds.
XX
OS Homo sapiens;
XX
PN M03946380-A2;
XX
PD 16-SEP-1999.
XX
PF 09-MAR-1999; 99MO-US05073.
XX
PR 13-MAR-1998; 98US-0039064.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Bandman O, Lai P, Hallman JL, Yue H, Corley NC;
XX Guggler KJ, Kaser MR, Baughn MR, Shah P;
XX
DR WPI; 1999-551409/46.
XX P-PSDB; AAY33298.
XX
PT MSP human membrane spanning proteins used to, e.g. prevent and treat
XX neoplastic disorders -
XX
PS Example 1; Page 80-81; 81pp; English.
XX
XX This invention describes novel human membrane spanning proteins (MSPs),
XX and the polynucleotides encoding them. The products of the invention are
XX used to diagnose, prevent and treat neoplastic, immunological and
XX reproductive disorders. This sequence encodes a human membrane spanning
XX protein MSP-5 fragment.
XX
SQ Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match 100.0%; Score 154; DB 20; Length 3111;
Best Local Similarity 100.0%; Pred. No. 1.4e-40;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGATCTTTTCATGCTATTCAGAGCTGTGATTAAGTCGATCGAATCAGAAATATCC 60
DB 2545 acagatcttttcattgctattcagcgtgtgataagtcgattcgaatcgaatatcc 2604
QY 61 AACATTGCAGAGATATCTTTGTTTATTCCTCCACAGACTCGCCAGACACCTAGTCT 120
DB 2605 aacattgcagagatattcttatttattcctccacagactcgcagagacacccagtcct 2664
QY 121 GATGAACGTCGTCTCTTCTCTTAATATTCATA 154
DB 2665 gatgaacgctctctcttcttccataatcata 2698

RESULT 9
ID AA229502 standard; cDNA; 3311 BP.
XX
AC AA229502;
XX

DT 12-Oct-2001 (first entry)
DE C902P determined cDNA sequence.
XX
XX Human; immunotherapy; diagnosis; colon cancer; colon tumour;
KW Immunogenic; gene therapy; vaccine; colonic cancer; ss.
OS Homo sapiens.
PN WO200149716-A2.
XX
XX 12-JUL-2001.
PD
XX
XX 29-DEC-2000; 2000WO-US35596.
XX
XX 30-DEC-1999; 99US-0476296.
PR 10-JAN-2000; 2000US-0480321.
PR 15-FEB-2000; 2000US-0504629.
PR 06-MAR-2000; 2000US-0519444.
PR 19-MAY-2000; 2000US-0575251.
PR 29-JUN-2000; 2000US-0609448.
PR 28-AUG-2000; 2000US-0649811.
XX
XX (CORI-) CORIXA CORP.
PA
XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX
XX WPI, 2001-441847/47.
DR
XX
XX Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer -
PS
XX Claim 2; Page 425-426; 472pp; English.

The present invention describes colon tumour associated proteins (I) and the polynucleotides (II) that encode them. (I) have cytostatic activity. (I) and (II) can be used in gene therapy and vaccine production. (I) and (II) may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate colon tumour associated protein (TCAP) expression, such as colonic cancer. For example, (I) and (II) may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of TCAPs by expressing inactive proteins or to supplement the patients own production of them. Additionally, (II) may be used to produce the TCAP proteins, by inserting the nucleic acids into a host cell culturing the cell to express the protein. (II) and its complementary sequences may also be used as DNA probes in diagnostic polymerase chain reaction (PCR) and hybridisation assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. (I) may also be used as antigens in the production of antibodies against TCAPs and in assays to identify modulators of TCAP expression and activity. Anti-(I) antibodies and antagonists may also be used to down regulate TCAP expression and activity. The anti-(I) antibodies may also be used as diagnostic agents for detecting the presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). AA128460 to AA129512 CC and AA124494 to AA125523 represent nucleotide and amino acid sequences given in the exemplification of the present invention.

Sequence 3311 BP; 1028 A; 692 G; 742 G; 849 T; 0 other;

Query Match 100.0%; Score 154; DB 22; Length 3311;
Best Local Similarity 100.0%; Pred.No. 1.4e+40;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 1 ACAGATCTTTCATTGCTCAATTCAGCGTGATGAAGGTGCATGGAATCAGAAATATCC 60
DB 2863 acgacatcttccgtcgtcatccagcgctgtgataaggctgcgatcgaatcaagaatatcc 2922
ACACTTCACGAGTAGTATCTTTGGTTTAATCTCCACAGACTCCGCCAGAGACACTAGTCC 120

```

D5 2523 aaccttgacagatgactcttggtttatcttcctccacagactccgcacagagaccctagctct 2982
QY 121 GATGAAAGCTGCTGCTCTCTTCTGCTTAATATTCATA 154
      |||||||
D5 2993 gatgaacgtctgctcttgccttccataatcata 3016
      |||||||

RESULT 110
AA145884
ID AA145884 standard; cDNA: 878 BP.
XX
XX AA145884;
XX
XX 13-MAR-1997 (first entry)
XX
XX Human colon specific gene CSG5 cDNA partial clone.
XX
XX Colon specific gene: CSG5; colon cancer; metastasis; diagnosis;
XX
XX gene therapy; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX
XX CDS 2..691
XX
XX FT /
XX
XX FT /*tag= a
XX
XX WC9639419-A1.
XX
XX PD 12-DEC-1996.
XX
XX PF 06-JUN-1995; 95WO-US07289.
XX
XX PR 06-JUN-1995; 95WO-US07289.
XX
XX PT (HUMA-) HUMAN GENOME SCI INC.
XX
XX RA
XX
XX RI Rosen CA, Yu G;
XX
XX DR WP1: 1997-043054/04.
XX
XX P-OSDB; AAM06548.
XX
XX
XX Human colon specific genes and their expression products - detection
XX
XX of which, in non-colon tissue samples, can be used as indication of
XX
XX colon cancer metastasis
XX
XX
XX Claim 1: Fig 5; 60pp: English.
XX
XX
XX 13. CDNA clones (AA145880-92), most of them partial clones, correspond
XX
XX to human colon specific genes, designated CSG1, CSG2, etc., that
XX
XX are primarily expressed in tissues derived from the colon. CSG7
XX
XX and CSG10 show reduced expression in colon cancer cells as compared
XX
XX to that in normal cells; the remaining genes are overexpressed in
XX
XX colon cancer. The partial cDNA sequences can be used to isolate
XX
XX full-length clones and genomic clones including the complete gene.
XX
XX CSG nucleic acids can be used to produce CSG polypeptides (see also
XX
XX AA06545-53) in transformed host cells, as probes to detect disorders
XX
XX of the colon, partic. colon cancer and colon cancer metastasis, and
XX
XX in gene therapy.
XX
XX
XX Sequence 878 BP; 257 A; 179 C; 188 G; 241 T; 13 other:

Query Match 92.9%; Score 143; DB 18; Length 878;
Best Local Similarity 99.4%; Pred. No. 3..5e-37;
Matches 154; Conservative 0; Mismatches 0; Indels 1; Gaps 1

QY 1 ACAGATCTTTTCAATTCGATATTCAGGCTGTTGATAGGTCGATCTGGAATACGAAATATCC 60
      |||||||
D5 545 acagatctttcattcgtctcttcagctgttgataagctcgatctgaaatcagaataatcc 604
      |||||||
QY 51 AACATTGCACGAGTATCTTTGTTTATTCGCCGACACACACGCCGACAGACACCTACTCTT 120
      |||||||
D5 605 aacattgcacgagatcttctgttattctctccacagactccgcacagagacctagctct 664
      |||||||

```


OY 121 GATGAACGTCGCTCTGT-CCTAATATTCATA 154
 DB 665 gatgaacgctcgtctctgtcctaataatcata 699

RESULT 11

AAVI6672
 ID AAVI6672 standard; cDNA; 878 BP.

AC AAVI6672;

DT 22-JUN-1998 (first entry)

DE Polynucleotide sequence of a colon-specific gene.

KW Colon-specific gene; probe; detection; expression; human;
 KW diagnostic assay; colon cancer; antibody; screening; ss.

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS 2..685
 FT /tag= a
 FT /note= "no stop codon given"

PN US5733748-A.

PD 31-MAR-1998.

PF 06-JUN-1995; 95US-0469667.

PR 06-JUN-1995; 95US-0469667.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen C, Yu G;

DR WPI; 1998-229823/20.

DR P-PSDB; AAW46879.

PT Colon-specific nucleic acids - useful as probes for detecting colon

PS cancer micrometastases

PS Claim 15; Fig 5A-B; 51pp; English.

CC AAVI6668-81 represent polynucleotide sequences of partial or full-length

CC cDNA clones of colon-specific genes. The polynucleotides can be used

CC as probes to detect expression of the corresponding human genes, e.g. in

CC diagnostic assays for detecting micrometastases of colon cancer.

CC Recombinant cells containing the polynucleotides can be used to

CC produce the polypeptides, in order that antibodies can be raised and

CC used in further screening or diagnostics.

CC Sequence 878 BP; 257 A; 179 C; 188 G; 241 T; 13 other;

CC Query Match 92.9%; Score 143; DB 19; Length 878;

CC Best Local Similarity 99.4%; Pred. No. 3.5e-37;

CC Matches 154; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 ACAGATCTTTTCATGCTATTCAGGCTGTGATTAAGTTCGATCGAATCAGAAATATTC 60
 DB 545 acagatcttttcattgctatcagcgtgtgataagtcgataatcagaataatc 604
 OY 61 AATATTCACAGAGATATCTTTGTTTATTCCTCCACAGACTCCGCCAGACACTAGTCT 120
 DB 605 aatattgcacgagatattctgttattcctccacagactccgcagagacactagct 664
 OY 121 GATGAACGTCGCTCTGT-CCTAATATTCATA 154
 DB 665 gatgaacgctcgtctctgtcctaataatcata 699

RESULT 12

AA722483
 ID AA722483 standard; cDNA to mRNA; 255 BP.

AC AA722483;

DT 22-AUG-1996 (first entry)

DE Human gene signature HUMGS04094.

KW gene signature; messenger RNA; mRNA; relative abundance; frequency;

KW human; cloning; mapping; non-biased library; diagnosis; detection;

KW cell typing; abnormal cell function; ss.

OS Homo sapiens

PN WO/514772-A1

PD 06-JUN-1995.

PF 11-NOV-1994; 94WO-JP01916.

PR 12-NOV-1993; 93JP-0355504.

PA (MATS/) MATSUBARA K.

PA (OKUB/) OKUBO K.

PI Matsubara K, Okubo K;

DR WPI; 1995-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.

PT for diagnosis of abnormal cell function, by preparing cDNA that

PT reflects relative abundance of corresp. mRNA in specific human

PT tissues

PS Claim 1; Page 1138-1139; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.

CC double-stranded DNA) which comprises one of the 7837 "GS" sequences

CC given in AAT19001-T26837 and which is able to hybridise to part of

CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)

CC sequences were obtained from 3'-directed cDNA libraries prepared

CC from various human tissues; synthesis of cDNA was initiated from the

CC 3' end of mRNA by using poly(T) as the sole primer. Since the 3'-

CC untranslated sequence is unique to a particular mRNA species, almost

CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library

CC is constructed so as to reflect accurately the relative abundance of

CC different mRNAs in the particular tissue from which it was derived.

CC The appearance frequency of a given GS in a cDNA library can be

CC determined (esp. using primers and probes derived from the GS

CC sequences) as a means of diagnosing abnormal cell function or for

CC recognising different cell types.

CC Sequence 255 BP; 76 A; 52 C; 39 G; 80 T; 8 other;

CC Query Match 72.7%; Score 112; DB 16; Length 255;

CC Best Local Similarity 97.4%; Pred. No. 3.1e-27;

CC Matches 112; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 10 GATCTGAATCAGAAATATTCACACATTGCACGAGTATCTTGTATTCCGCCACAGACT 99
 DB 101 gatctgaatnagaataatccaacattgcacgagatcttnttatnccccaagact 60
 OY 100 CCGCAGAGACACCTAGTCTGTGATGAACGTCGCTCTGTCTTAATATTCATA 154
 DB 101 ccgcagagacacctagctctgtatgaacgctcgtccttgcctaataatcata 115
 RESULT 13
 AAH46101

ID	AAH46101 standard; DNA: 2739 BP.
XX	AAH46101;
AC	11-SEP-2001 (first entry)
XX	Mouse Gob-5 coding sequence, SEQ ID NO:3.
XX	Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
XX	KW expression inhibition; antisense therapy; gene therapy; bronchial asthma;
KW	chronic obstructive pulmonary disease; antiasthmatic; ds.
XX	
OS	Mus sp.
XX	
XX	Key Location/Qualifiers
XX	CD5 1..2739
FT	/*tag= a
FT	/partial
FT	/product= "Mouse Gob-5"
FT	/note= "No stop codon given in the specification"
XX	WO200138530-A1.
XX	31-MAY-2001.
XX	22-NOV-2000; 2000WO-JP08232.
XX	24-NOV-1999; 99JP-0333479.
XX	27-APR-2000; 2000JP-0127589.
XX	(TAKE) TAKEDA CHEM IND LTD.
XX	Nakanishi A, Morita S;
XX	WPI: 2001-355935/37.
XX	P-PSDB: AAB73715.
XX	New antisense nucleotide, useful for treatment and prevention of
XX	bronchial asthma and chronic obstructive pulmonary disease -
XX	
XX	Claim 3; Page 80-82; 104pp: Japanese.
XX	
XX	The invention relates to an antisense nucleotide targetted to the mouse
XX	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX	the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX	relates to an antibody specific for the Gob-5 protein, medical and
XX	diagnostic compositions containing the antisense nucleotide or the
XX	antibody, and methods and kits for screening for compounds which inhibit
XX	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX	The antisense oligonucleotides and antibody are therefore useful for the
XX	treatment and prevention of bronchial asthma and chronic obstructive
XX	pulmonary disease. The present sequence represents the mouse Gob-5
XX	gene coding sequence.
XX	
XX	Sequence 2739 BP: 784 A; 687 C; 655 G; 613 T; 0 other:
XX	
XX	Query Match 47.8%; Score 73.6; DB 22; Length 2739;
XX	Best Local Similarity 81.7%; Pred. No. 2.6e-14;
XX	Matches 85; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX	
XX	1 ACAGATCTTTTCAATCTATTCAGGCTGTGATTAAGTGCATCTGAAATCAGAAATATTC 60
XX	2512 acagatctttctctctcctcccaagctgtgtaagtccaatctgaatcagaatctctc 2571
XX	61 AACATTGACGAGATATCTTTGTTATTCCTCCACAGACTCGCC 104
XX	2572 aacattgcacgggtgctgtgtcatcccccgcctcagagacgcgc 2615
XX	
XX	RESULT 14
XX	AAH46120 standard; cDNA: 2843 BP.

XX AAH46120;
XX
XX 11-SEP-2001 (first entry)
XX
XX
DE Mouse Gob-5 cDNA, SEQ ID NO:22.
XX
XX
XX Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
XX expression inhibition; antisense therapy; gene therapy; bronchial asthma;
XX chronic obstructive pulmonary disease; antiasthmatic; ss.
XX
XX Mus sp.
XX
XX Location/Qualifiers
FH Key 15..2756
FH CDS /*tag= a
FH /product= "Mouse Gob-5"
XX
XX W0200138530-A1.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-JP08232.
XX
XX 24-NOV-1999; 99JP-0333479.
PR 27-APR-2000; 2000JP-0127589.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
PI Nakanishi A, Morita S;
PI
TR WIT. 2001-355935/37.
DR P-RSDB; AAB73715.
XX
XX New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease -
XX
XX Example 1; Page 89-91; 104pp; Japanese.
XX
XX The invention relates to an antisense nucleotide targetted to the mouse
XX Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX relates to an antibody specific for the Gob-5 protein, medical and
XX diagnostic compositions containing the antisense nucleotide or the
XX antibody, and methods and kits for screening for compounds which inhibit
XX the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX The antisense oligonucleotides and antibody are therefore useful for the
XX treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
XX
XX Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;

Query Match 47.8%; Score 73.6; DB 22; Length 2843;
Best local Similarity 81.7%; Pred. No. 2,6e-14;
Matches 85; Conservative 0; Mismatches 19; Indels 0; Gaps 0.

CY 1 ACAGATCTTTTCATTGCTCATTCAGCGCTGTGAAGAAGTGCATCTGAAACAGAAATATCC 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2586 acagatatcttcattgcctaccacgagcgtcgatgaagtccatcttgaaataacagaaatctcc 2585

CY 61 AACATTGCAGAGTAGTATCTTGTTTATCTCCTCCACAGACTCCGCC 104
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2586 aacattgcacgggtgctgtgttcatcccgcgtcacgagagcgcc 2629

RESULT 115
AAAF81925
ID AAFF81925 standard; CDNA; 2931 BP.
XX
XX AAFF81925;
XX
XX
DT 13-JUN-2001 (first entry)

Search Completed: April 4, 2002, 03:18:25

Job time: 60257 sec

```

XX DE Murine ICACC-1 nucleotide sequence.
XX KM ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
XX KM Interleukin 9 induced calcium activated chloride channel; IL-9;
XX KM calcium activated chloride channel; anti-allergic; anti-asthmatic;
XX KM anti-inflammatory; immunomodulatory; cystic fibrosis;
XX KM inflammatory bowel disease; autoimmune disease; ss.
XX OS Mus sp.
XX FH Key Location/Qualifiers
XX FT 8.2749
XX FT /tag= a
XX FT /product= "ICACC-1"
XX FT /note= "IL-9 induced calcium activated chloride channel"
XX PN W09944620-A1.
XX PD 10-SEP-1999.
XX PF 03-MAR-1999; 99MO-US04703.
XX PR 03-MAR-1998; 98US-0076815.
XX PA (MAGA-) MAGAININ PHARM INC.
XX PI Holroyd KJ, Levitt RC, Maloy WL, Louahed J, McLane M;
XX PI Nicolalde NC, Zhou Y, Dong Q;
XX DR WPI: 1999-550979/46.
XX DR P-PSDB: AAB74822.
XX PT New nucleic acid encoding calcium activated chloride channel, used to
XX PT identify, e.g. specific modulators for treating atopic allergy -
XX PS Claim 2; Fig 2; 75pp; English.
XX CC The present sequence encodes the murine interleukin 9 (IL-9) induced
XX CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
XX CC have anti-allergic, anti-asthmatic, anti-inflammatory and
XX CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
XX CC used to alleviate asthma (or more generally atopic allergy), while those
XX CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
XX CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
XX CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma
XX CC (increased levels) or IBD (reduced levels), also for monitoring
XX CC treatment of these conditions. The ICACC proteins can be used:
XX CC (1) to raise specific antibodies (Ab), useful: (a) as immunoassay
XX CC reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
XX CC (A) to disrupt bonding between ICACC and its ligands; and (iii) to
XX CC identify modulators and binding partners. ICACC polynucleotides can be
XX CC used to generate transgenic animals or recombinant cells, used to screen
XX CC for antagonists, also as a source of therapeutic antisense agents or
XX CC diagnostic probes (for quantifying mRNA expression, e.g. for
XX CC identification of modulators).
XX SQ Sequence 2931 BP; 867 A; 719 C; 684 G; 661 T; 0 other;

Query Match 47.8%; Score 73.6; DB 20; Length 2931;
Best Local Similarity 81.7%; Pred. No. 2,6e-14;
Matches 85; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ACAGATCTTTCATGCTTCAGAGCTGTGATAGTCGATCGAATCAGAAATATCC 60
db 2519 acagatcttcattcattcagcgtcgataaagtcacattcgaataatcc 2578

QY 61 AACATTGACGAGATCTTTGTTATTCCTCCACAGACTCCGCC 104
db 2579 aacattgcacgggtgtctgtcattcccccgcgtcagagcgcgc 2622

```


Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:28:45 : Search time 16681 Seconds
(without alignments)
1921.623 Million cell updates/sec

Title: US-09-049-696-20

Perfect score: 2983

Sequence: 1 GAATCAGAGGAGATGTAC.....AAATGCTAACACTGGGTA 2983

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 segs, 537289281 residues 22703874

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Lasting first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_est1:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1760.8	59.0	2933	12 AK008659	AK008659 Mus muscu
2	1746.6	58.6	2915	12 AK007466	AK007466 Mus muscu
3	692.8	23.2	712	10 AM009763	AM009763 ws87b04.x
4	652.4	21.9	716	10 A1660234	A1660234 ws68g02.x
5	648.4	20.5	734	10 A1660957	A1660957 wf20d08.x
6	612.6	20.5	653	10 AM361534	AM361534 QV2-CT026
7	567.8	19.0	598	11 BE927204	BE927204 RCI-CT024
8	565	18.9	633	10 AM753451	AM753451 QV2-CT026
9	557	18.7	625	10 AM361532	AM361532 QV2-CT026
10	553	18.5	657	10 AM361521	AM361521 QV2-CT026
11	551.4	18.5	621	10 AM361520	AM361520 QV2-CT026
12	551.2	18.5	561	10 A1802756	A1802756 wf18a04.x

13	520	17.4	652	10 AM753449	AM753449 QV2-CT026
14	516.8	17.3	528	11 BF083270	BF083270 RCI-CT024
15	489	16.4	504	10 A1802693	A1802693 wf17c01.x
16	487.2	16.3	561	10 AM361523	AM361523 QV2-CT026
17	480	16.1	501	10 A1721275	A1721275 as82h08.x
18	473.6	15.9	480	10 AA581198	AA581198 nd38c07.x
19	461.8	15.5	654	10 AM361522	AM361522 QV2-CT026
20	450	15.1	490	10 AA508854	AA508854 n122h10.s
21	445	14.9	450	10 AM009764	AM009764 ws87b05.x
22	444	14.9	460	10 AM750670	AM750670 CM2-CM03
23	439.4	14.7	806	11 BG962080	BG962080 602826824
24	437.2	14.7	466	11 BG195179	BG195179 RST14358
25	434	14.5	455	11 BF083321	BF083321 RCI-CT024
26	432.4	14.5	496	10 AM753448	AM753448 QV2-CT026
27	421.8	14.1	661	11 BF581041	BF581041 602100173
28	421.8	14.1	935	11 BF578919	BF578919 602095492
29	411.8	13.8	906	11 BF581342	BF581342 602100726
30	399	13.4	767	11 BF578246	BF578246 602093073
31	397.6	13.3	704	11 BG965023	BG965023 602831668
32	391.6	13.1	745	11 BG963539	BG963539 602831217
33	384.6	12.9	890	11 BF578833	BF578833 602094876
34	379.4	12.7	911	11 BF578693	BF578693 602093175
35	378.8	12.7	876	11 BF234977	BF234977 602027529
36	369.4	12.4	974	11 BG962464	BG962464 602827322
37	369.2	12.4	998	11 BF583412	BF583412 602097421
38	368	12.3	392	10 A1687981	A1687981 wa77f06.x
39	362.6	12.2	812	11 BG968878	BG968878 602836186
40	358.8	12.0	435	10 AM844161	AM844161 RCI-CT026
41	352.8	11.8	356	10 AM361528	AM361528 QV2-CT026
42	352.4	11.8	378	10 AA296955	AA296955 EST112726
43	349.8	11.7	747	11 BF580957	BF580957 602100659
44	336.8	11.3	646	10 AA691335	AA691335 vs14h06.x
45	336	11.3	732	11 BG963719	BG963719 602831354

ALIGNMENTS

RESULT 1	AK008659	2933 bp	mRNA	HTC	05-JUL-2001
LOCUS	AK008659		Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210006g11, full insert sequence.		
DEFINITION	AK008659.1	GI:12842987			
ACCESSION	AK008659				
VERSION					
KEYWORDS	CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 2933)				
TITLE	Carninci, P. and Hayashizaki, Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Methods in enzymology. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	10349636				
AUTHORS	2 (bases 1 to 2933)				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to				
MEDLINE	Prepare full-length cDNA libraries for rapid discovery of new genes				
PUBMED	Genome research. 10 (10), 1617-1630 (2000)				
REFERENCE	2049374				
AUTHORS	3 (bases 1 to 2933)				
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, A., Sumi, N., Ishi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,				

Db 809 CCAATGACCAAAACCAAGATGCAATCTCCGAGACAGCTGGGAATGCATCCAGGAATCT 868
Oy 862 GAGGACTTTAAGAAAACCACTCTATGACAACAAGCACCACCAATCCACACTTCTCATTTG 921
Db 869 GAGGACTTCAAGCAACCACTCCATGAGAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 928
Oy 922 CTGCAATTTGACAAAGAAATTTGTGTGTTTATGCTTGTGACAAATCTGGAAAGCAGTGC 961
Db 929 CTGCAATTTGACAAAGAAATTTGTGTGTTTATGCTTGTGATTAAGTCCGGAGACATGCTG 988
Oy 982 GGTAAAGGCTGCAATGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1041
Db 989 GATGATCTGCTTAAACGAATGAATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1048
Oy 1042 CTGGGGCTCTGGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1101
Db 1049 CAGGATCTCTGGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1108
Oy 1102 ATACAGATTAACAGTGGACAGTGGACAGGAGACACCTGCCAAAGATTAACCTGACAGCT 1161
Db 1109 AAACAGTTAAACAGTGGTGTGACAGAGATCTGATCAAGCATTACCCACAGATATCT 1168
Oy 1162 TCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1221
Db 1169 GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1228
Oy 1222 TATCCAACTGATGATCTGAAATTTGTGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 1281
Db 1229 TATCCAACTGATGATCTGAAATTTGTGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 1288
Oy 1282 GGTGCTTTTACGAGGTCACAAAGAGTGGTCCATATCCACAGTGGTGGGGGCC 1341
Db 1289 AGTGTGCTTACGAGGTCACAAAGAGTGGTCCATATCCATATGAGTGGGGGCC 1348
Oy 1342 TCTGAGCTCAAGAACTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1401
Db 1349 GCTGCGGCTTAAAGAGCTTGAAGAGCTGTGCAAAATATACAGAGGAGGAGGAGGAGGAGGAG 1408
Oy 1402 TCAGATCAAGTTCAGAAACATGAGCTCATTTGATGATGATGATGATGATGATGATGATGATG 1461
Db 1409 TCGATCAAGTTCAGAAACATGAGCTCATTTGATGATGATGATGATGATGATGATGATGATG 1468
Oy 1462 GGAAGTCTCTCAGGCTTCATCCAGCTTGAAGATTAACCTCCAGAAAGC 1521
Db 1469 GCGGCGATGCTCAGGCTTCATCCAGCTTGAAGATTAACCTCCAGAAAGC 1528
Oy 1522 CAGTGTGAATGAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1581
Db 1529 CAGTGTGAATGAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1588
Oy 1582 ATACCTGAGCAACGAGGCTCCCAAAATCTCTCTGAGATCCAGTGGAGAGAGCA 1641
Db 1589 ATACCTGAGCAACGAGGCTCCCAAAATCTCTCTGAGATCCAGTGGAGAGAGCA 1648
Oy 1642 GGTGCTTTTGTAGTGGACAAACCAAAAGGCTTACCTCCAAATCCAGGCAATTTGCT 1701
Db 1649 AAGGCTTTTGTAGTGGACAAACCAAAAGGCTTACCTCCAAATCCAGGCAATTTGCT 1708
Oy 1702 AAGGCTTTTGTAGTGGACAAACCAAAAGGCTTACCTCCAAATCCAGGCAATTTGCT 1761
Db 1709 AAGGCTTTTGTAGTGGACAAACCAAAAGGCTTACCTCCAAATCCAGGCAATTTGCT 1768
Oy 1762 AGGTCCCGTGGCTCAATGCTGCTGCTCAATTAAGTGAAGTTCACAAAGCAAG 1821
Db 1769 ACCTCCCGTGGCTCAATGCTGCTGCTCAATTAAGTGAAGTTCACAAAGCAAG 1828
Oy 1822 GACACGAGCAAAATCCCGAGGCTCTGCTGATTAATTAATTTGAGGAGAGGAGGAGGAGGAG 1881
Db 1829 AACACGAGCAAAATCCCGAGGCTCTGCTGATTAATTAATTTGAGGAGAGGAGGAGGAGGAG 1888
Oy 1882 CCAATTTCTAGGAGGAGTGCAGAGGCTGATTAATTAATTTGAGGAGAGGAGGAGGAGGAG 1941
Db 1889 CCAATTTCTAGGAGGAGTGCAGAGGCTGATTAATTAATTTGAGGAGAGGAGGAGGAGGAG 1948

Oy 1942 TTGGAACCTACTGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2001
Db 1949 CTGGAATTAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2008
Oy 2002 AGGTATTTTCAACCTTAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2061
Db 2009 AGGTATTTTCAACCTTAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2068
Oy 2062 GAGATTAAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2121
Db 2069 GAGATTAAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2128
Oy 2122 GGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2181
Db 2129 GGTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2188
Oy 2192 GATGTTTCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2241
Db 2193 TATGTTTCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2242
Oy 2242 TCTGATGCTC---CAATGCTCCATACCTGATCTCTCCACCTGAGGAGGAGGAGGAGGAGGAG 2298
Db 2243 ACCAATGCTCCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2302
Oy 2299 CTGAGGCGGAGAAATTCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2358
Db 2303 CTGAGGCGGAGAAATTCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2362
Oy 2339 GATTATGACCTAGGAACAGCTCAACAGATATATCAATTCAGGAGGAGGAGGAGGAGGAGGAGGAG 2418
Db 2353 GAGTATGACCTAGGAACAGCTCAACAGATATATCAATTCAGGAGGAGGAGGAGGAGGAGGAGGAG 2422
Oy 2419 CTGAG 2478
Db 2423 CTGAG 2482
Oy 2479 GCCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2538
Db 2483 GCGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2542
Oy 2539 GATCTTTTCAATCTTATTCAGGCTGTGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2598
Db 2543 GATCTTTTCAATCTTATTCAGGCTGTGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2602
Oy 2599 ATTGACGAGTATCTTGTATTTATTCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2658
Db 2603 ATTGACGAGTATCTTGTATTTATTCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2653
Oy 2659 GAAAGCTGCTCTCTGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2718
Db 2654 GAAAGCTGCTCTCTGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2713
Oy 2719 TTTAAATTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2778
Db 2714 CTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2773
Oy 2779 TTTGTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2838
Db 2774 TTTGTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2825
Oy 2839 TGTATTTAGACTTCTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2874
Db 2836 TGTATTTAGACTTCTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2861
RESULT: 2
AK007466 2915 bp mRNA HTC 05-JUL-2001
LOCUS: Mus musculus 10 day old male pancreas cdna, RIKEN full-length
DEFINITION: enriched library, clone:1810012P03, full insert sequence.
ACCESSION: AK007466
VERSION: AK007466.1 GI:12841032

Oy 802 CCNAAACAAATTCNAAAAATGCAATCTCCGACAGCATGGGAATGATCCGTGATTCT 861
 Db 809 CCAAAATGACCAAAACCAAGATGATCTCCGACAGCATGGGAATGATCCGTGATTCT 868
 Oy 862 GAGGACTTAAAGAAAACCACTCTATGACACAGACCCAAATCCGACCTTCTGATG 921
 Db 869 GAGGACTTAAAGAAAACCACTCTATGACACAGACCCAAATCCGACCTTCTGATG 928
 Oy 922 CTGACAGATGGACAAAGAAATGCTGTTAGTCTTGGACAAATCTGGAACTGGCGACT 981
 Db 929 CTGCAAAATTTGACAAAGAAATGCTGTTAGTCTTGGAAATCTGGGAGCAATGGCGAAC 988
 Oy 982 GGTAAACCCCTTAATGACTGATCAAGACAGCCAGCTTTCTGCTGACAGACATTGAG 1041
 Db 989 GATGATCTGCTTGGACCAATGATGATGACCAAGCCGCTTTCTGCTGACAGCTGGAG 1048
 Oy 1042 CTGGGCTCTGGGTTGGGATGGTGAATTTGACAGTGGCCCATGCTACAAAGTGAAC 1101
 Db 1049 CAGGATCTGGGTTGGGATGGTGAATTTGACAGTGGCCCATGCTACAAAGCGAATCT 1108
 Oy 1102 ATACAGATTAACAGTGGACAGTGGACAGGACACACTGCGCAAAAAGATTACCTGACAGCT 1161
 Db 1109 AAACAGTTAAACAGTGGTCTGACAGAGATCTGCTGATCAAGACACTTACCAACAGTATCT 1168
 Oy 1162 TCAGAGGAGACCTCCATCTGACAGCGGCTTGCATGGCAATTACTGTGATTAAGAAAGA 1221
 Db 1169 CGAGAGGAGACATCTATATGCTCTGCGCTTGGACAGCATTTTACAGTATTAAGAAAGA 1228
 Oy 1222 TATTCAGATGATGATCTGAATTTGCTGCTGACAGATGGGGAAGACAACTATAGT 1281
 Db 1229 TATTCAGATGATGATCTGAATTTGCTGCTGACAGATGGGGAAGACAACTATAGT 1288
 Oy 1282 GGGTCTTTAAACGAGTCAAAACAAAGTGGTGCATTCATCAACAGATCTGTTGGGCGCC 1341
 Db 1289 ACTGCTTTGACCTGGTGAAGACAGCGGGCGCTCATCAATACATGGCCCGGAGACCG 1348
 Oy 1342 TCTGACGTCAGAACTAGAGAGCTGTCCAAATATGACAGAGG -TTTACAGATATGCT 1400
 Db 1349 GCTGCCGCTAAAGAGCTTGAAGAGCTGTCCAAATATGACAGAGGCTCTGACAGATATGCT 1408
 Oy 1401 TTCAGATCAAGTTCAGAAACAAATGGCTCATTTGCTTGGGCGCTTTCATCAAGAA 1460
 Db 1409 TTTGATCTGCTGACAAACAAATGGCTCATTTGCTTGGGCGCTTTCATCAAGAA 1468
 Oy 1461 TGGACCTGCTCTGACAGCGCTCCATCCAGCTTGAAGTAAGCTTAAACCTGACAAAG 1520
 Db 1469 TCGCGGAGATCGCTCAGCACTCCATCCAGCTGAGAGACAGGGAGTTAACTCCAGAAATA 1528
 Oy 1521 CCAGTGGATGAATGGACAGTATGCTGAGACAGCCGTGGGAAGGACACTTGTCTCT 1580
 Db 1529 CCAATGGATGAATGGCTCAGTATGCTGAGACAGCTGCTGGGCAAGGACACCTTGTCTCT 1588
 Oy 1581 TATTCAGTGGACAAAGCAAGCCCTCCCAATTCCTCTCTGGGATCCCAAGGACAGAAAGA 1640
 Db 1588 TATTCAGTGGACAAAGCAAGCCCTCCCAATTCCTCTCTGGGATCCCAAGGAGGTGAAGA 1648
 Oy 1641 AGGTGGCTTTTGAAGTGAAGAAACCAAAATGGCTTCACTTCAATCCAGAGCAATTC 1700
 Db 1649 AAATGGTTTAACTAGACAAACCACTAAAGTGGCTTCACTTCAATCCAGAGCGAG 1708
 Oy 1701 TTAAGTGGCACTTGGAAATTAAGTCTGACAGCAAGCTTCAAACTTGAACCTGACTGT 1760
 Db 1709 TTAAGTGGCTTTTGAAGTGAAGTGAAGTCAAGTCAAGACCTTCACTTGAACCTGACTGT 1768
 Oy 1761 CAGGTCCGCTGGTCCAAATGCTACCTGCTCAATTTAGAGTACTTCAAAAGCAAGA 1820
 Db 1769 CAGCTCCGCTGGTCCAAATGCTACCTGCTCAATTTAGAGTACTTCAAAAGCAAGA 1828
 Oy 1821 GGACACAGCAAAATTTCCAGGCTCTCTGTAATTTATGCAAAATATTTCCGCAAGAGACT 1880
 Db 1829 GAAACAGAGGAAATTTCCAGGCTCTCTGTAATTTATGCAAAATATTTCCGCAAGAGACT 1888
 Oy 1881 CCCAATTTCTCAGGGCCAGTGTCAAGCCCTGATTGAATCAAGTAATGAAAAACAGTTAC 1940

Db 1899 GCTATTTCTCAGGGCCAGGCTCACACCTTGAATTTGTAATGTGTAAGAAAAACAGTAAC 1948
 Oy 1941 CTGGAAGTACTGATTAAGAGAGAGAGTCTGATGCTACTTAAAGATGAGAGGCTGCTACT 2000
 Db 1949 CTTGGAATTTACTGATTAAGAGAGAGAGTCTGATGCTACTTAAAGATGAGAGGCTGCTACT 2008
 Oy 2001 AAGTATTTTCAACACTTATGACAGCAATGATGATTAAGTAAAGTGGGCTCTGAG 2060
 Db 2009 AAGTATTTTCAACACTTATGATGCAAAATGATGATGATGATGATGATGATGATGATGAT 2068
 Oy 2061 AAGGATTAAGAGAGAGAGAGAGAGTATACCCAGAGAGTGGAGACTGTACTATAC 2120
 Db 2069 AAGGATTAAG 2128
 Oy 2121 TGGCTGATTTGAGATGATGATTAAGTAAATGATGATGATGATGATGATGATGATGAT 2180
 Db 2129 TGGCTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2182
 Oy 2181 TGAATTTCAACAGCAAGTGTGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2240
 Db 2183 TGAATTTCAACAGCAAGTGTGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2242
 Oy 2241 TTTGATGCTC---CAATGCTCCATTAATGCTTCTTCCACCTTGGCCAAATCAAG 2297
 Db 2243 CACCAATGCTCCCGAG 2302
 Oy 2298 CCGTGAAGGCGAAATTCACAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2357
 Db 2303 CCGTGAAGGCGAG 2362
 Oy 2368 TGAATTTCAACAGCAAGTGTGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2417
 Db 2369 TGAATTTCAACAGCAAGTGTGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2422
 Oy 2418 TCTCAG 2477
 Db 2423 TCTCAG 2482
 Oy 2478 AGCAG 2537
 Db 2483 AGCAG 2542
 Oy 2538 AGATCTTTTCAATGATTAAG 2597
 Db 2543 AGATCTTTTCAATGATTAAG 2602
 Oy 2598 CATTCAG 2657
 Db 2603 CATTCAG 2653
 Oy 2658 TGAAG 2717
 Db 2654 AGATCTTTTCAATGATTAAG 2713
 Oy 2718 TTTAAATATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2777
 Db 2714 GGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2773
 Oy 2778 TTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2837
 Db 2774 TTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2825
 Oy 2838 ATGATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2874
 Db 2826 ATGATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2862

RESULT 3
 AM009763/C
 LOCUS
 DEFINITION
 AM009763 712 bp mRNA EST 08-MAR-2000
 ws87b04.x1 NCL-CCAP_C03 Homo sapiens cDNA IMAGE:2504911 3'
 similar to TR:088826 088826 GDB-5 PROTEIN ; mRNA sequence.


```

Db 260 CAGGCTGTGATGATGCTGATCTGAATTCAGAAATATCCAAATTCAGACGATCTTTG 201
OY 2617 TTTATTCCTCAGACACTCCGCCAGACACCTAGCTCTGCTGTAAGACCTGCTGCTTGT 2676
Db 200 TTTATTCCTCAGACACTCCGCCAGACACCTAGCTCTGCTGTAAGACCTGCTGCTTGT 141
OY 2677 CCTATATTCATATCAACAGACACCTCTGCGATTCATCATATTTTAAATTTATGTGGAAG 2736
Db 140 CCTATATTCATATCAACAGACACCTCTGCGATTCATCATATTTTAAATTTATGTGGAAG 81
OY 2737 TGAATGAGAGAACTGACGCTGTCAATAGCCTTGAATTTTTCAGATTAATATAA 2796
Db 80 TGAATGAGAGAACTGACGCTGTCAATAGCCTTGAATTTTTCAGATTAATATAA 21
OY 2797 TAAATCATTCATCTTTT 2814
Db 20 TAAATCATTCATCTTTT 3

```

```

RESULT 6
LOCUS AM361534 653 bp mRNA EST 04-FEB-2000
DEFINITION QV2-CT0261-261099-011-F01 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361534
VERSION AM361534.1 GI:6866288
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

```

REFERENCE 1 (bases 1 to 653)
AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?pl=QV2&t2=QV2-CT0261-
261099-011-F01&t3=1999-10-26&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 37
High quality sequence stop: 652.
Location/Qualifiers

```

```

FEATURES
source
1..653
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/notes="Organ: colon; Vector: puc18; Site: 1: Small, site: 2:
Small; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. letters patent application No. 196
'716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 140 a 156 c 166 g 191 t
ORIGIN

```

```

Query Match 20.5%; Score 612.6; DB 10; Length 653;
Best Local Similarity 99.1%; Pred. No. 3e-134;
Matches 637; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

```

OY 1517 ACAGCCAGTGA- GAATGSCACAGTGAATGTCGACACACCGCTGGGAAAGACACTTGT 1576
|||||

```

```

Db 644 ACAGCCAGTGA- GAATGSCACAGTGAATGTCGACACACCGCTGGGAAAGACACTTGT 586
OY 1577 TTTATTCCTCAGACACTCCGCCAGACACCTAGCTCTGCTGTAAGACCTGCTGCTTGT 1636
Db 1585 TTTATTCCTCAGACACTCCGCCAGACACCTAGCTCTGCTGTAAGACCTGCTGCTTGT 526
OY 1637 AGCAAGTGGCTTTGAGTGAACAAAACACCAAAATGAGCTTACCTCAATCCAGGCA 1696
Db 525 AGCAAGTGGCTTTGAGTGAACAAAACACCAAAATGAGCTTACCTCAATCCAGGCA 466
OY 1697 TTGCTAAGTTGGGACCTTGAATATGACGCTGTGACAGCAAGCTTACCAATCCAGGCA 1756
Db 465 TTGCTAAGTTGGGACCTTGAATATGACGCTGTGACAGCAAGCTTACCAATCCAGGCA 406
OY 1757 CTGTGACGCTCCCTGGCTGCTCAATGCTACCTGCTCAATTTACAGTACTTCCAAACGA 1816
Db 405 CTGTGACGCTCCCTGGCTGCTCAATGCTACCTGCTCAATTTACAGTACTTCCAAACGA 346
OY 1817 ACAAGGACACACCAAAATTCGCCAGCCCTGTGTAATTTATGCAAAATTTCCGCAAGAG 1876
Db 345 ACAAGGACACACCAAAATTCGCCAGCCCTGTGTAATTTATGCAAAATTTCCGCAAGAG 286
OY 1877 CTGTGACGCTCCCTGGCTGCTCAATGCTACCTGCTCAATTTACAGTACTTCCAAACGA 1936
Db 295 CTGTGACGCTCCCTGGCTGCTCAATGCTACCTGCTCAATTTACAGTACTTCCAAACGA 226
OY 1937 TTACCTTGAACACTGATGATATGACAGGCTGCTGATGCTACTAAGATGACGCTGTCT 1996
Db 225 TTACCTTGAACACTGATGATATGACAGGCTGCTGATGCTACTAAGATGACGCTGTCT 166
OY 1997 ACTGAGCTATTTACAACTTATGACACGAATGCTATGATGATGATGATGATGATGATGAT 2056
Db 165 ACTGAGCTATTTACAACTTATGACACGAATGCTATGATGATGATGATGATGATGATGATGAT 106
OY 2057 TGGGAGGAGTTAAAGCAGACCGCAGAGAGATGATACCCAGCAGAGTGAAGTGAAGTGAAGTGA 2116
Db 105 TGGGAGGAGTTAAAGCAGACCGCAGAGAGATGATACCCAGCAGAGTGAAGTGAAGTGAAGTGA 46
OY 2137 TACCTGG-CTGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2158
Db 45 TACCTGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3

```

```

RESULT 7
LOCUS BE927204 586 bp mRNA EST 02-OCT-2000
DEFINITION RCI-CT0249-240800-211-h10 CT0249 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE927204
VERSION BE927204.1 GI:10453280
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

```

REFERENCE 1 (bases 1 to 586)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Negal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202653
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

```

```

TITLE
JOURNAL
MEDLINE
COMMENT:

```

```

Query Match 20.5%; Score 612.6; DB 10; Length 653;
Best Local Similarity 99.1%; Pred. No. 3e-134;
Matches 637; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```


QY 1795 ATTACAGTACTTCCAAAAGAACAGACACGCAATTTCCCGCCCTGTGAGTT 1854
 DB 312 ATTACAGTACTTCCAAAAGAACAGACACGCAATTTCCCGCCCTGTGAGTT 253
 QY 1955 TATGCAATATTTGCCAAGAGAGCTCCCAATTTCTCAGGGCCAGTGTACAGCCCTGAT 1914
 DB 252 TATGCAATATTTGCCAAGAGAGCTCCCAATTTCTCAGGGCCAGTGTACAGCCCTGAT 193
 QY 1915 GAATCAGTAAATGAAAAACAGTTTACCTTGAAGTACTGATATGAGAGAGTGCAT 1974
 DB 192 GAATCAGTAAATGAAAAACAGTTTACCTTGAAGTACTGATATGAGAGAGTGCAT 193
 QY 1975 GCTACTAGAGTGAAGGCTGTCTACTCAAGTATTTCAACTTATGACAGAGTGTAGA 2034
 DB 132 GCTACTAGAGTGAAGGCTGTCTACTCAAGTATTTCAACTTATGACAGAGTGTAGA 73
 QY 2035 TACACTGTAAAGTGGGCTCTGGAGAGTAAAGCAGCCAGAGAGAGTATGCC 2094
 DB 72 TACACTGTAAAGAGGCTCTGGAGAGTAAAGCAGCCAGAGAGAGTATGCC 13
 QY 2095 CAGCA 2099
 DB 12 CAGCA 8
 RESULT 9
 AM361532 625 bp mRNA EST 04-FEB-2000
 LOCUS QV2-CT0261-261099-011-e03 CT0261 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 AM361532
 VERSION
 AM361532.1 GI:6866286
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 HCGP http://www.ludwig.org.br/ORESTES.
 The FAPESP/LICR Human Cancer Genome Project
 Unpublished (1999)
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV2-CT0261-
 261099-011-e03&tl3=1999-10-26&tl4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 93
 Location/Qualifiers
 1..625
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0261"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
 Sma1; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 7716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 172 a 154 c 150 g 149 t
 ORIGIN

Query Match 18.7% Score 557 DB 10; Length 625;
 Best Local Similarity 99.18; Pred. No. 4.2e-121;
 Matches 560; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1235 GATTCGAATATTTGCTGCTGAGAGATGGGGAAGAACACACTATTAAGTGGCTTTAAGC 1294
 DB 61 GATTCGTAATTTGCTGCTGAGAGATGGGGAAGAACACACTATTAAGTGGCTTTAAGC 120
 QY 1295 AGGTCAAAACAAAGTGTGCATCATCCACACAGTCGCTTTGGGGCCCTCAGCTCAAG 1354
 DB 121 AGGTCAAAACAAAGTGTGTGCTTCATGACAGTGTGCTTTGGGGCCCTCAGCTCAAG 180
 QY 1355 AACTGAGAGAGTGTCCAAATATGACAGAGTGTACAGACATATGCTTCAGATCAAGTTC 1414
 DB 131 AACTGAGAGAGTGTCCAAATATGACAGAGTGTACAGACATATGCTTCAGATCAAGTTC 240
 QY 1415 AGAACAATGGCCCTCATTTGATGCTTTGGGGCCCTTCATCAGGAATATGAGCTGTCTC 1474
 DB 231 AGAACAATGGCCCTCATTTGATGCTTTGGGGCCCTTCATCAGGAATATGAGCTGTCTC 300
 QY 1435 AGCGCTCCATCCAGCTTGTGAGAGTAAAGGATTAACCTCCAGAACAGCCAGTGAATG 1534
 DB 301 AGCGCTCCATCCAGCTTGTGAGAGTAAAGGATTAACCTCCAGAACAGCCAGTGAATG 360
 QY 1535 GCACAGTATGCTGAGACAGACCGTGGAAAGACACTTTGTTCTATCAGCTGAGACA 1594
 DB 351 GCACAGTATGCTGAGACAGACCGTGGAAAGACACTTTGTTCTATCAGCTGAGACA 420
 QY 1595 CGCAGCTCCCTCCAAATCTCTCTGGAGTCCAGTGGACAGAAAGAGTGGCTTTGAG 1654
 DB 431 CGCAGCTCCCTCCAAATCTCTCTGGAGTCCAGTGGACAGAAAGAGTGGCTTTGAG 480
 QY 1635 TGGCAAAAACACCAAAATATGCTTACCTCCAAATCCAGGCAATGCTTAAGTGGCACTT 1714
 DB 431 TGGCAAAAACACCAAAATATGCTTACCTCCAAATCCAGGCAATGCTTAAGTGGCACTT 540
 QY 1715 GGAATATCAGTGTGACAGCAAGTCAAAACCTTGACCTGACGTGACGCTGCGCT 1774
 DB 541 GGAATATCAGTGTGACAGCAAGTCAAAACCTTGACCTGACGTGACGCTGCGCT 600
 QY 1775 CCAATGCAACCTGCTCCCAATTAC 1799
 DB 601 CCAATGCTACCTGCTCCCAATTAC 625
 RESULT 10
 AM361521 657 bp mRNA EST 04-FEB-2000
 LOCUS QV2-CT0261-261099-011-f03 CT0261 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 AM361521
 VERSION
 AM361521.1 GI:6866275
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 HCGP http://www.ludwig.org.br/ORESTES.
 The FAPESP/LICR Human Cancer Genome Project
 Unpublished (1999)
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV2-CT0261-
 261099-011-f03&tl3=1999-10-26&tl4=1)
 Seq primer: puc 18 forward

High quality sequence start: 123
High quality sequence stop: 656.
Location/Qualifiers

FEATURES

source

1. 657
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 175 a 160 c 164 g 158 t
ORIGIN

Query Match 18.5%; Score 553; DB 10; Length 657;
Best Local Similarity 97.3%; Pred. No. 3.8e-120;
Matches 573; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 1198 GCATTACTGATATAGAGAAATATCACTGATGGA-TCTGAATGTCTGCTGAC 1256

DB 69 GCATTAAATGATAGAGAAATATCACTGATGGA-TCTGAATGTCTGCTGAC 128

QY 1257 GCATGGGAGAGACACATATATAGTGGGCTTTAAAGAGGCAAGGTCGACAT 1316

DB 129 GCATGGGAGAGACACATATATAGTGGGCTTTAAAGAGGCAAGGTCGACAT 188

QY 1317 CATCCACACAGTGGCTTTGGGGCCCTGACAGCTCAAGAACTAGAGAGCTGCCAAT 1376

DB 189 CATCCACACAGTGGCTTTGGGGCCCTGACAGCTCAAGAACTAGAGAGCTGCCAAT 248

QY 1377 GACAGAGGTTAGACATATGCTTGAATGCAATGCAATGCAATGCAATGCAATG 1436

DB 249 GACAGAGGTTAGACATATGCTTGAATGCAATGCAATGCAATGCAATGCAATG 308

QY 1437 TTTTGGGGCCCTTTGATCAGAAATGAGAGCTGCTCAGCCCTCCATCCAGCTTGAGAG 1496

DB 309 TTTTGGGGCCCTTTGATCAGAAATGAGAGCTGCTCAGCCCTCCATCCAGCTTGAGAG 368

QY 1437 TAAGGATTAACCTCCAGACAGCAGTGGATGAATGGACAGATGCTGGACAGAC 1556

DB 369 TAAGGATTAACCTCCAGACAGCAGTGGATGAATGGACAGATGCTGGACAGAC 428

QY 1557 CGTGGGAAGACACTTTGTTCTTATACCTGAGCAACAGCAGCTCCCAATCTCTT 1616

DB 429 CGTGGGAAGACACTTTGTTCTTATACCTGAGCAACAGCAGCTCCCAATCTCTT 488

QY 1617 CTGGATCCCAAGTGGACAGAAAGAGGCTTTGATGAGCAAAACACCAAAATGAG 1676

DB 489 CTGGATCCCAAGTGGACAGAAAGAGGCTTTGATGAGCAAAACACCAAAATGAG 548

QY 1677 CTACCTCCAAATCCAGGATGCTAAGGTTGGACCTTGGAAATACAGTCTCCAAAG 1736

DB 549 CTACCTCCAAATCCAGGATGCTAAGGTTGGACCTTGGAAATACAGTCTCCAAAG 608

QY 1737 CTCAAAACCTGACCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 1785

DB 609 CTCAAAACCTGACCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 657

RESULT 11
LOCUS AM361520 621 bp mRNA EST 04-FEB-2000
DEFINITION OY2-CT0261-261099-011-e07 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361520
VERSION AM361520.1 GI:6866274
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 621)
HCCP <http://www.ludwig.org.br/ORESTES>.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?pl1-QY2<2-QY2-CT0261-261099-011-e07<3-1999-10-26<4-1>)
Seq primer: puc 18 forward
High quality sequence start: 88
High quality sequence stop: 613.
Location/Qualifiers

FEATURES

source

1. 621
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 166 a 151 c 150 g 154 t
ORIGIN

Query Match 18.5%; Score 551.4; DB 10; Length 621;
Best Local Similarity 96.2%; Pred. No. 9e-120;
Matches 586; Conservative 0; Mismatches 21; Indels 2; Gaps 2;

QY 1137 GCGTTCGATCGGATTTACTGATATAGAGAAATATCACTGATGGA-TCTGAATGTCTGCTGAC 1246

DB 11 GCGTTCGATCGGATTTACTGATATAGAGAAATATCACTGATGGA-TCTGAATGTCTGCTGAC 69

QY 1247 TCGTGTGACGATGGGGAAGACACATATATAGTGGGCTTTAAAGAGGTCGACGCTCAACAA 1306

DB 70 TCGTGTGACGATGGGGAAGACACATATATAGTGGGCTTTAAAGAGGTCGACGCTCAACAA 129

QY 1307 GTGGTCCATATCCACAGACAGTGGGCTTTGAGGCGCTGACGCTCAAGAACTAGAG-GAG 1365

DB 130 GTGGTCCATATCCACAGACAGTGGGCTTTGAGGCGCTGACGCTCAAGAACTAGAG-GAG 189

QY 1396 CTGTCCAAATGACAGAGGTTTACAGACATATGTTGATCAATGATCAATGATCAATGAT 1425

DB 190 CTGTCCAAATGACAGAGGTTTACAGACATATGTTGATCAATGATCAATGATCAATGAT 249

QY 1426 CTCAATGATGCTTTGGGGCCCTTTGATCAAGGAAATGAGAGCTGCTCAGGCTCAGC 1485

DB 250 CTCAATGATGCTTTGGGGCCCTTTGATCAAGGAAATGAGAGCTGCTCAGGCTCAGC 309

QY 1486 CAGCTTGAGAGTAAGGATTAACCTCCAGAACAGCAGTGGATGATGATGATGATGATG 1545

DB 310 CAGCTTGAGAGTAAGGATTAACCTCCAGAACAGCAGTGGATGATGATGATGATGATG 369

QY 1546 GTGGACAGCAGCTGGGAAAGACACTTTGTTCTTATACCTGAGCAACGACGCTCC 1605

DB 370 GTGGACAGCAGCTGGGAAAGACACTTTGTTCTTATACCTGAGCAACGACGCTCC 429

QY 1696 CAAATCCTTCTCGGGATCCAGTGACAGCAAGAGGCTGCTTACTGACCAAAAC 1665

DB 430 CAAATCCTTCTCGGGATCCAGTGACAGCAAGAGGCTGCTTACTGACCAAAAC 489

FEATURES	Location/Qualifiers
source	1. .652

BASE COUNT	179 a	153 c	161 g	159 t
ORIGIN				

BASE COUNT	158 a	113 c	127 g	130 t
ORIGIN				

Db 361 TGAAGCGGAAATTCACGGGGGCGAGTCTCATTAATCTGACTTGACACAGCTCCTGGGGATG 420

QY 2360 ATTATGACCATGAGACAGCTCAGCAATATATCATTTGCAATAGTACAGTATTTTGTATC 2419

Db 421 ATTATGACCATGAGACAGCTCAGCAATATATCATTTGCAATAGTACAGTATTTTGTATC 480

QY 2420 TCAGAGCAGAGTTCATGAAATCTCTCAAGTGAATCTACTGCTCTCA 2467

Db 481 TCAGAGCAGAGTTCATGAAATCTCTCAAGTGAATCTACTGCTCTCA 528

RESULT 15
AI802693 504 bp mRNA EST 18-DEC-1999
LOCUS AI802693/c
DEFINITION wfl7c01.x1 Soares Dieckgraefe.colon_NHUC Homo sapiens cDNA clone
IMAGE:2350848 3' similar to YR:088826 088826 GDB-5 PROTEIN. ;, mRNA
sequence.

ACCESSION AI802693 GI:5368165
VERSION AI802693
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 504)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 727 Std Error: 0.00
Seq primer: -40UP from Gidco
High quality sequence stop: 392.

FEATURES
Source Location/Qualifiers
1..504
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2350848"
/clone_lib="Soares_Dieckgraefe.colon_NHUC"
/tissue_type="colonic mucosa from 5 ulcerative colitis
patients"
/lab_host="pDH10B (phage-resistant)"
/note="Organ: colon; Vector: pUT73D-Pac (Pharmacia) with a
modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAGAGTGGAGCGCGCTAGCTTTTCTTTTCTTTT 3'],
TGTACCAATCTGAGAGTGGAGCGCGCTAGCTTTTCTTTTCTTTT 3'],
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pUT73 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraefe (Washington University,
dieck@im.wustl.edu); colonic mucosa represents a range of
disease involvement from mild cryptitis to severe
ulceration, fibrosis, and degeneration. Library
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 150 a 85 c 103 g 163 t 3 others
ORIGIN

Query Match 16.4%; Score 489; DB 10; Length 504;
Best Local Similarity 99.4%; Pred. No. 5e-105;
Matches 489; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2323 AGTCTCATTAATCTGACTTGACAGCTCTCGGGAGATGATGACATGAGACAGCTCAC 2382

Db 504 AGTCTCATTAATCTGACTTGACAGCTCTCGGGAGATGATGACATGAGACAGCTCAC 445

QY 2383 AAGTATATCATTCGAAATAGACAGATATCTTGATCTCAGAGACAAATTCATGATCT 2442

Db 444 AAGTATATCATTCGAAATAGACAGATATCTTGATCTCAGAGACAAATTCATGATCT 385

QY 2443 CTTCAAGTGAATCTACTGCTCTCATATCCCAAGGAGCCAACTCTGAGAGCTTTTGG 2502

Db 334 CTTCAAGTGAATCTACTGCTCTCATATCCCAAGGAGCCAACTCTGAGAGCTTTTGG 325

QY 2503 TTTAAGCAGAAACATTAATCTTTGAAATGGCAGACATCTTTTCATTTGCTATTCAGGCT 2562

Db 334 TTTAAGCAGAAACATTAATCTTTGAAATGGCAGACATCTTTTCATTTGCTATTCAGGCT 265

QY 2563 GTTCATAGTGTGATCTGAAATTCAGAAATATCCCAACATTCAGAGATATCTTTGTTTAT 2622

Db 264 GTTCATAGTGTGATCTGAAATTCAGAAATATCCCAACATTCAGAGATATCTTTGTTTAT 205

QY 2682 COTCCAGAGCTCGGCGCAGAGACACCTAGCTGATGAAAGCTGCTCTTGTCTAAT 2682

Db 145 COTCCAGAGCTCGGCGCAGAGACACCTAGCTGATGAAAGCTGCTCTTGTCTAAT 145

QY 2693 ATTCAATATCAACAGCAGCAATTCCTGCGCATTCACATTTTAAATTTAGGAGTGGATA 2742

Db 144 ATTCAATATCAACAGCAGCAATTCCTGCGCATTCACATTTTAAATTTAGGAGTGGATA 85

QY 2743 GGAGAACTGACAGCTGTCATATAGCCTAGGCGTGAATTTTGTGAGATAATTAATAATC 2802

Db 94 GGAGAACTGACAGCTGTCATATAGCCTAGGCGTGAATTTTGTGAGATAATTAATAATC 25

QY 2803 ATTCAATCTTTT 2814

Db 24 ATTCAATCTTTT 13

Search completed: April 3, 2002, 20:29:25
Job time: 40328 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:56:42 ; Search time 521.98 Seconds
(without alignments)
1294.272 Million cell updates/sec

Title: US-09-049-696-20

Perfect score: 2983
Sequence: 1 GAATTCACAGGAGATGTAC.....AAATGCTAACCACTGGGTA 2983

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	790.8	26.5	878	1	US-08-469-667-8
2	790.8	26.5	878	5	PCT-US95-07289-8
3	790.8	10.9	401	4	US-09-221-298-34
4	228.2	7.7	576	4	US-09-385-982-23
5	221.4	7.4	595	4	US-09-385-982-25
6	200.8	6.7	618	4	US-09-385-982-24
7	183.4	6.1	611	4	US-09-385-982-27
8	168.6	5.7	742	4	US-09-385-982-33
9	52	1.7	7218	1	US-08-232-463-14
10	38.6	1.3	5156	1	US-09-091-432-3
11	36.4	1.2	8700	2	US-08-392-625-16
12	36.4	1.2	8700	2	US-08-466-961A-16
13	36.4	1.2	8700	2	US-08-645-193B-18
14	36.4	1.2	8700	2	US-08-464-517-36
15	36	1.2	3158	2	US-08-246-361A-36
16	36	1.2	3158	4	US-08-463-772-36
17	35.2	1.2	3058	3	US-09-276-531-107
18	35.2	1.2	1794	3	US-09-012-515A-13
19	35	1.2	1794	3	US-08-360-144A-13
20	35	1.2	1794	5	PCT-US95-06722-13
21	34.8	1.2	805	1	US-08-118-469A-6
22	34.8	1.2	805	1	US-08-909-119-6
23	34.8	1.2	6924	1	US-08-015-973-2
24	34.8	1.2	6924	1	US-08-448-164-2
25	34.8	1.2	19124	2	US-08-487-826B-13
26	34.6	1.1	5923	4	US-09-064-922-3
27	34	1.1	3600	4	US-08-855-910-7

28	34	1.1	5319	1	US-08-169-927-1	Sequence 1, Appl
29	33.8	1.1	1566	4	US-08-871-572B-2	Sequence 2, Appl
30	33.8	1.1	2255	4	US-08-871-572B-3	Sequence 3, Appl
31	33.8	1.1	3447	1	US-08-252-995D-3	Sequence 3, Appl
32	33.8	1.1	3447	2	US-08-834-108-3	Sequence 3, Appl
33	33.6	1.1	3183	1	US-08-413-135-1	Sequence 1, Appl
34	33.6	1.1	4090	4	US-08-569-214-4	Sequence 4, Appl
35	33.6	1.1	4090	4	US-08-937-236-4	Sequence 4, Appl
36	33.6	1.1	4810	3	US-08-852-629-11	Sequence 11, Appl
37	33.2	1.1	2520	2	US-08-454-557C-50	Sequence 50, Appl
38	33.2	1.1	2520	2	US-08-340-426D-50	Sequence 50, Appl
39	33.2	1.1	2520	2	US-08-450-673C-50	Sequence 50, Appl
40	33.2	1.1	2520	5	PCT-US95-17111A-50	Sequence 50, Appl
41	33.2	1.1	3095	6	5231168-1	Sequence 50, Appl
42	33.2	1.1	10684	3	US-08-618-100B-3	Sequence 3, Appl
43	33	1.1	1984	1	US-07-885-970A-25	Sequence 25, Appl
44	33	1.1	1985	1	US-08-298-687A-25	Sequence 25, Appl
45	33	1.1	1985	1	US-08-298-829-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-08-469-667-8
Sequence 8, Application US/08469667
Patent No. 5733748
GENERAL INFORMATION:
APPLICANT: YU, Guo-Liang
APPLICANT: Rosen, Craig
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carrella, Byrne, Bain, Gillfillan, Cecchi,
Stewart & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,667
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-435
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..665
US-08-469-667-8

Query Match 26.5%; Score 790.8; DB 1; Length 878;
Best Local Similarity 97.9%; Pred. No. 4,6e-225;
Matches 820; Conservative 1; Mismatches 15; Indels 2; Gaps 2;

QY 1992 TGTCTACTCAAGGATATTTCACAACTATGACACCAATGTAAGATACAGTGTAAAGTCCG 2051
 DB 1 TGTCTACTCAAGGATATTTCACAACTATGACACCAATGTAAGATACAGTGTAAAGTCCG 60
 QY 2052 GGCCTCTGGAGAGTAAAGCAGCAGCAGAGAGATGATACCCAGAGAGATGGAGACT 2111
 DB 61 GGCCTCTGGAGAGTAAAGCAGCAGCAGAGAGATGATACCCAGAGAGATGGAGACT 120
 QY 2112 GTACATACCTGGCTGGATGAGATGAAATGAAATACATGAAATCCACAGACTGAAAT 2171
 DB 121 GTACATACCTGGCTGGATGAGATGAAATGAAATACATGAAATCCACAGACTGAAAT 180
 QY 2172 TAATAGATGATGTTTCAACACAGCAAGTGTGTTTCAAGCAAGATCTCGGGAGGCTC 2231
 DB 181 TAATAGATGATGTTTCAACACAGCAAGTGTGTTTCAAGCAAGATCTCGGGAGGCTC 240
 QY 2232 ATTTGGCTGTGATGTCACCAATGCTCCATACCTGATCTCTCCAGCTGGCCAAAT 2291
 DB 241 ATTTGGCTGTGATGTCACCAATGCTCCATACCTGATCTCTCCAGCTGGCCAAAT 300
 QY 2292 CACCGACCTGAAAGCGGAAATTCACGGGGGCGAGTCTCATTAATCTGACTGGACAGCTCC 2351
 DB 301 CACCGACCTGAAAGCGGAAATTCACGGGGGCGAGTCTCATTAATCTGACTGGACAGCTCC 360
 QY 2352 TGGGATGATTTTGAACATGAAAGCTCACAAGTATATCATTCGATTAAGTACAGATAT 2411
 DB 361 TGGGATGATTTTGAACATGAAAGCTCACAAGTATATCATTCGATTAAGTACAGATAT 420
 QY 2412 TCTTGATCTCAGAGACAACTGATGAAATCTCTCAAGTAACTACTGCTCATCC 2471
 DB 421 TCTTGATCTCAGAGACAACTGATGAAATCTCTCAAGTAACTACTGCTCATCC 480
 QY 2472 AAAGGACCAACTCTGAGAGAGTCTTTTGTGTTTAAACAGAAACATTAATCTTTGAAA 2531
 DB 481 AAAGGACCAACTCTGAGAGAGTCTTTTGTGTTTAAACAGAAACATTAATCTTTGAAA 540
 QY 2532 TGGCAGAGATCTTTCATGCTATTCAGGCTGTGATTAAGTGTGATCTGAAATCAGAAAT 2591
 DB 541 TGGCAGAGATCTTTCATGCTATTCAGGCTGTGATTAAGTGTGATCTGAAATCAGAAAT 600
 QY 2592 ATCCACATGTCAGAGAGATCTTTCATGCTATTCAGGCTGTGATTAAGTGTGATCTGAAAT 2651
 DB 601 ATCCACATGTCAGAGAGATCTTTCATGCTATTCAGGCTGTGATTAAGTGTGATCTGAAAT 660
 QY 2652 TCCGTGATGAAAGCTGCTGCTCTGTGATTAAGTGTGATCTGAAATCAGAAAT 2710
 DB 661 TCCGTGATGAAAGCTGCTGCTCTGTGATTAAGTGTGATCTGAAATCAGAAAT 720
 QY 2711 TTCACATTTTAAATTAATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2770
 DB 721 TTCACATTTTAAATTAATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 780
 QY 2771 GCTGATTTTTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2828
 DB 781 GCTGATTTTTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 837

RESULT 2
 PCT-US95-07289-8
 Sequence 8, Application PC/TUS9507289

GENERAL INFORMATION:

APPLICANT: Yu, Guo-Liang

APPLICANT: Rosen, Craig

TITLE OF INVENTION: Colon Specific Genes and Proteins

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carella, Byrne, Bain, Giffillan, Cecchi,

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: NJ

COUNTRY: USA

ZIP: 07068-1739

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/07289
 FILING DATE: 06-JUN-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Ferraro, Gregory D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 325800-265
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 878 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 2..685
 PCT-US95-07289-8

Query Match 26.5%; Score 790.8; DB 5; Length 878;
 Best Local Similarity 97.9%; Pred. No. 4.0e-225;
 Matches 820; Conservative 1; Mismatches 15; Indels 2; Gaps 2;

QY 1992 TGTCTACTCAAGGATATTTCACAACTATGACACCAATGTAAGATACAGTGTAAAGTCCG 2051
 DB 1 TGTCTACTCAAGGATATTTCACAACTATGACACCAATGTAAGATACAGTGTAAAGTCCG 60
 QY 2052 GGCCTCTGGAGAGTAAAGCAGCAGCAGAGAGATGATACCCAGAGAGATGGAGACT 2111
 DB 61 GGCCTCTGGAGAGTAAAGCAGCAGCAGAGAGATGATACCCAGAGAGATGGAGACT 120
 QY 2112 GTACATACCTGGCTGGATGAGATGAAATGAAATACATGAAATCCACAGACTGAAAT 2171
 DB 121 GTACATACCTGGCTGGATGAGATGAAATGAAATACATGAAATCCACAGACTGAAAT 180
 QY 2172 TAATAGATGATGTTTCAACACAGCAAGTGTGTTTCAAGCAAGATCTCGGGAGGCTC 2231
 DB 181 TAATAGATGATGTTTCAACACAGCAAGTGTGTTTCAAGCAAGATCTCGGGAGGCTC 240
 QY 2232 ATTTGGCTGTGATGTCACCAATGCTCCATACCTGATCTCTCCAGCTGGCCAAAT 2291
 DB 241 ATTTGGCTGTGATGTCACCAATGCTCCATACCTGATCTCTCCAGCTGGCCAAAT 300
 QY 2292 CACCGACCTGAAAGCGGAAATTCACGGGGGCGAGTCTCATTAATCTGACTGGACAGCTCC 2351
 DB 301 CACCGACCTGAAAGCGGAAATTCACGGGGGCGAGTCTCATTAATCTGACTGGACAGCTCC 360
 QY 2352 TGGGATGATTTTGAACATGAAAGCTCACAAGTATATCATTCGATTAAGTACAGATAT 2411
 DB 361 TGGGATGATTTTGAACATGAAAGCTCACAAGTATATCATTCGATTAAGTACAGATAT 420
 QY 2412 TCTTGATCTCAGAGACAACTGATGAAATCTCTCAAGTAACTACTGCTCATCC 2471
 DB 421 TCTTGATCTCAGAGACAACTGATGAAATCTCTCAAGTAACTACTGCTCATCC 480
 QY 2472 AAAGGACCAACTCTGAGAGAGTCTTTTGTGTTTAAACAGAAACATTAATCTTTGAAA 2531
 DB 481 AAAGGACCAACTCTGAGAGAGTCTTTTGTGTTTAAACAGAAACATTAATCTTTGAAA 540
 QY 2532 TGGCAGAGATCTTTCATGCTATTCAGGCTGTGATTAAGTGTGATCTGAAATCAGAAAT 2591
 DB 541 TGGCAGAGATCTTTCATGCTATTCAGGCTGTGATTAAGTGTGATCTGAAATCAGAAAT 600

Accession	Sequence	Position
QY 2592	ATCCACATTGCACGAGTATCTTTGTTTATTCCTCCACAGACTCCGCCAGAGACACCTTACG	2651
Db 601	ATCCACATTGCACGAGTATCTTTGTTTATTCCTCCACAGACTCCGCCAGAGACACCTTACG	660
QY 2652	TCTGTATGAAACGCTCTGCTCTCTGTGCTCTATATTTATATCAACAGACACATCTCTGGCA	2710
Db 661	TCTGTATGAAACGCTCTGCTCTCTGTGCTCTATATTTATATCAACAGACACATCTCTGGCA	720
QY 2711	TTTCACATTTTAAAAATTATGTGCAAGTGCATAGGAGACATGCACGCTGTCAATAGCCTTAG	2770
Db 721	TTTCACATTTTAAAAATTATGTGCAAGTGCATAGGAGACATGCACGCTGTCAATAGCCTTAG	780
QY 2771	GCTGAATTTTTCGACGATTAATAAATAATCAATATCTTTTTCGATATATATAA	2828
Db 781	GCTGAATTTTTCGCTGGTGATTAATATATSAATTCATCCTTTTTTTCGTTATATATAA	837

```

RESULT      3
US-09-221-298-34
: Sequence 34, Application US/09221298
: Patent No 6284241
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
: TITLE OF INVENTION: OF COLON CANCER
: FILE REFERENCE: 210121.471
: CURRENT APPLICATION NUMBER: US/09/221,298
: CURRENT FILING DATE: 1998-12-23
: NUMBER OF SEQ. ID NOS: 112
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ. ID NO 34
: LENGTH: 401
:
: TYPE: DNA
: ORGANISM: Human
:
: US-09-221-298-34

```

Query Match	10.9%	Score 323.8	DB 4	Length 401
Best Local Similarity	96.7%	Pred. No. 1.7e-86		
Matches 384; Conservative	0	Mismatches 7	Indels 6	Gaps 5

109 AACATGGGTATGAAAGGCACTTTCGTGGAATCAGACCCCAATGTGGCAGAAATGAAACA 168
 Db 1 aacaatgggtatgaaggcatctgcttgcaatcgaccaccaatgctgcagaatgaaaca 60
 169 CTCATTCAACAATAAAGACATGTGGACCCAGGCATCTCTATCTGTTCAAGCTACA 228
 Db 61 ccattcaacaacaataaagacatgctgaccceagcactcctgtatctgtttgaagtaca 120
 229 GGAAAGCATTTTATTTCAAAAATGTGCCATTTTGATTCCTGAACAATGAGACAAAG 288
 Db 121 ggaagaacgatttattatccaaatgctgcatcttgcattccctgaacaatgagaacaag 180
 289 GGTGACTATGTGAGACCAAACTTGAGACCTMCMAAAAATGCTGATCTTCTGTGTCG-TCGA 347
 Db 181 gcttgactaagtgaagaccaaaacttgagaccctcaaaaatgctgatatctgtgtgtcttga 240
 348 GTCCTACTCTCCAGATTAATGATGAACCCACACTGAGCAGAT-GGGCACTGTGAGAGA 406
 Db 241 gctactactcctcaagtaataatgatagaacctcaactgagcagatggggcaactcttgagagaa 300
 407 AGGG--TGAAAGGATCCACCTCCTCACTCTGATTTCATTTGACGAAAAAAAGTTAGC-TGAAT 463
 Db 301 agggggtgaagaagatccaccactcactcctctgattctcatcttgacgagaaaagttagcttgaat 360
 464 ATGAGCACCAAGGT-AGGCGATTTGTCCATGAGTGGG 499
 Db 361 atgagccacaagtaaggacatcttgccatgtaaltgg 397

RESULT 4
US-09-385-982-23
: Sequence 23, Application US/09385982

Patent No.65262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TYPE OF INVENTION: PRODUCTS: II
FILE REFERENCE: CCDDA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(576)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-23

```

: Query Match 7.7%; Score 228.2; DB 4; Length 576;
: Best Local Similarity 69.1%; Pred. No. 5.3e-58;
: Matches 385; Conservative 163; Indels 9; Gaps 5

```

QY	1882	CCATTTCGACGGGCGCACTGTGCACAGCCCTGATTGAATCATGTAATGAAAAACAGTTACC	194
DB	112	cccttctcttgagacaatgagactgctcttcattgaatcacagaattgacatcacagaagt	61
QY	1942	TTGAACTACTGGATTAATGAGAGGGGCGATGCTACTAAAGATGAGCGCTGTACTGCA	200
DB	62	tttgaaccttttgataatggtgcagcgcgcgactcttccaagaatgagatggaaatctacc	121
QY	2002	AGGTAATTTCACCAACTTATGACACAGCAATGTGATATACAGTGTAAAGTGGCGCTCTGGGA	206
DB	132	agatattttaaagacatatagaaatgagatatagcttaaaagttcggctcatgga	181
QY	2062	GGAACTTAACGACGACGACAGCGAAGATGATACCCACAGTGGAGCACTGTACTACT	212
DB	132	ggagcaaacacatcgcagcgcttaaaatlaacggtccatgaatagagtcgctgatacca	241
QY	2122	GGCGTGAATTAGATGATGTAATCAATGAATGAATCCACAAAGCCTGAATTAATGAAGAT	218
DB	212	ggctcggtgtagtgaacgggggaattgaagacaaccgcgaagaccgaaattgat---	298
QY	2182	GATGTTCAACACAAAGCAAGTGTGTTTCAGACAGATTCCTCGGAGGCTCATTTGTGCT	224
DB	239	gatactcagacacaccttgagggatttcagccgacaagcatcggaggtcatttgta	358
QY	2242	TCTATATCCCAATGTGCCATCACTGATCTTCCACACTGGCCAAATCACCACTG	230
DB	359	tcaaaatgcccaagc-cttctctgctgacccaataccacaagtaaatcacagaact	417
QY	2302	AAGCGGAATTAATCACGGGGGCGACTGTCAATTAATCTACTTGGACAGCCTCTGGGGTGAT	236
DB	418	gagtcacacagttcatttgg---ataagattatcttcaatctgacacacacagagataat	474
QY	2352	TATGACCAATGGAACAGCTCCCAAGTATATCATTTGCAATTAAGTACATATTTCTTGATCTC	242
DB	475	tttagtcttggaanaagttcaacgttatatcataagaataa-tgcgcagttcttga-cta	532
QY	2432	AGAGACAAGTTCAATGA	2438
DB	593	agagacagtttgatga	549

RESULT 5

CURRENT APPLICATION NUMBER: US/09/385, 982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328, 111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117, 393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098, 639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 27
LENGTH: 611
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(611)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-27

Query Match 6.1%; Score 183.4; DB 4; Length 611;
Best Local Similarity 68.0%; Pred. No. 1.1e-44;
Matches 300; Conservative 0; Mismatches 136; Indels 5; Gaps 3;

QY 1882 CCAATTCTCAGGCGCAGTGTACAGCCCTGATTGAATCAGTGAAT- GGAATAACAGTTAC 1940
DB 2 cctgttcttgagcccaattgactgtcttcaatgaaatcagaatgagatcagaagt 61
QY 1941 CTTGGACACTGCTGATTAATGAGCAGAGTGTGCTACTAAGAGATGACGCTTACTC 2000
DB 62 ttgagactcttgagaaatgagcagcgtgacttcttcaagaatgagatgagctc 121
QY 2001 AAGGTATTTCACAACTTATGACAGATGATGATAGTAAAGTGGGCTCGG 2060
DB 122 caggtatcttaccagcatatacagaataatgagataatagcttaaaagctcgctcag 181
QY 2061 AGGATTTAAGCAGCAGCAGCAGAGAGTGTATACCCAGCAGAGTGTGACATATAC 2120
DB 182 agagagaacaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 241
QY 2121 TGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2180
DB 242 aagcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 298
QY 2181 TGATGTTCACACAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2240
DB 299 gatactacgaccaccttgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 358
QY 2241 TTC-TGATGTTCACAAATGCTCCATACCTGATGATGATGATGATGATGATGATGATGAT 2299
DB 359 atcacaagaatcccaaaccttctcctgctcagcagcagcagcagcagcagcagcagcagc 418
QY 2300 TGAAGCGGGAATTCACGGGG 2320
DB 419 ttgatgtccacaagctcattag 439

RESULT 8
US-09-385-982-33
Sequence 33, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS: II
FILE REFERENCE: CCDA-260XX
CURRENT APPLICATION NUMBER: US/09/385, 982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328, 111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117, 393
EARLIER FILING DATE: 1999-01-27

EARLIER APPLICATION NUMBER: 60/098, 639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 33
LENGTH: 742
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(742)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-33

Query Match 5.7%; Score 168.6; DB 4; Length 742;
Best Local Similarity 69.9%; Pred. No. 3.2e-40;
Matches 285; Conservative 0; Mismatches 115; Indels 8; Gaps 4;

QY 2174 ACATACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2173
DB 111 acatcacagccttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 60
QY 2174 ATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2233
DB 61 atagagatc---tcagaccacttgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 117
QY 2234 TTGTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2293
DB 118 ttgagatcacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 177
QY 2294 CCAGCAGGAGCGGGAATTCACGCGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2353
DB 178 cagaccttgatgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 234
QY 2354 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2413
DB 285 gataatcttgatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 294
QY 2414 TTGATGTCACAGCAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2473
DB 295 ttgatcagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 352
QY 2474 AGGAGCGCAACTGTCAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2521
DB 353 agagagc 400

RESULT 9
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:

Thu Apr 4 09:27:41 2002

us-09-049-696-20.in

Page 6

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZqpt-Fls
US-08-232-463-14

Query Match 1.7%; Score 52; DB 1; Length 7218;
Best Local Similarity 4.9%; Pred. No. 5.5e-05;
Matches 19; Conservative 21; Mismatches 156; Indels 0; Gaps 0;

OY 442 GCAGGAAAAAGTGTGATGATGACACACAGGAGGCAATTTGTCATGAGTGGGCT 501
DB 1463 GTAGTTAAAGATAGACAAATTTGTCACRRRRRRRRRRRRRRRRRRRRRRRR 1404
OY 502 CATCTAGATGGAGCTATTGACGAGTAAATGATGAGAAATTTACTTAATCAAT 561
DB 1403 RRR 1344
OY 562 GCAGAGTACAGCAGTATGTTGACGAGATATGATGTAATGTAAGTAAGAG 621
DB 1343 RRR 1284
OY 622 TGTACAGGAGGAGCTGTACACAAAGATGCATTCATTAAGTAAGAGCTTAT 681
DB 1283 RRR 1224
OY 682 GAAAAAGATGTGATTTGTTCCATCCCGCAGACGAGAGGCTCTAATGTT 741
DB 1223 RRR 1164
OY 742 GCACAAGATGTGATCTATGTAATCTGTACAGAAACCAACCAAGAGAGCT 801
DB 1163 RRR 1104
OY 802 CCAACAGCAAAATCAAAATGCAA 827
DB 1103 RRR 1078

RESULT 10
US-09-091-432-3/C
Sequence 3, Application US/09091432
Patent No. 5981837
GENERAL INFORMATION:
APPLICANT: Chapelle, Clint
TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
FILE REFERENCE: 7024-325
CURRENT APPLICATION NUMBER: US/09/091,432
CURRENT FILING DATE: 1998-06-18
EARLIER APPLICATION NUMBER: PCT/US96/20094
EARLIER FILING DATE: 1996-12-19
EARLIER APPLICATION NUMBER: US 60/009,119
EARLIER FILING DATE: 1995-12-22
EARLIER APPLICATION NUMBER: US 60/013,388

EARLIER FILING DATE: 1996-03-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Microsoft Word 2.0C
SEQ ID NO 3
LENGTH: 5156
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-091-432-3

Query Match 1.3%; Score 38.6; DB 2; Length 5156;
Best Local Similarity 50.8%; Pred. No. 0.43;
Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

OY 647 AAGATGCATTCATTAAGTAAAGAGGACTCTATGAAAAAGATGTGATGTTGCTCC 706
DB 4835 AAAAGGTGAATTAACAAAAAATTTTCTTTTAAAGTAGAAGTTCTATCC 4776
OY 707 AATCCCGCCAGAGGAGAGGCTTCTAATGTTGACACACATGTTGATTCATAGTTG 766
DB 4775 AAGATGCATCTTTGACATGCTAATTAATTAATGAAAAAATCTATTACTGTTA 4716
OY 767 AATCTGTACAGAAACCAACCAAGAGCTCCAAACGCAAAATCAAAATGCA 826
DB 4715 CATTCATTCAGAGAAATCCCATTAAGAAAAAAGGCAACAGAGACCA 4656
OY 827 A 827
DB 4655 A 4655

RESULT 11
US-08-392-625-16/C
Sequence 16, Application US/08392625
Patent No. 5837485
GENERAL INFORMATION:
APPLICANT: Enliam, Karl-Dieter
APPLICANT: Gtz, Friedrich
APPLICANT: Schnell, No. 5837485bert
APPLICANT: Augustin, Johannes
APPLICANT: Engelke, Gernar
APPLICANT: Rosensteln, Ralf
APPLICANT: Kaletta, Cortina
APPLICANT: Klein, Cora
APPLICANT: Wleland, Bernd
APPLICANT: Kupke, Thomas
APPLICANT: Jung, G nther
APPLICANT: Kellner, Roland
TITLE OF INVENTION: Of Chemical Compounds
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,625
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/876,791
FILING DATE: 30-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893

Thu Apr 4 09:27:41 2002

DB 1279 TCTAGTCTCTCTATTAGACATAAATACCTGTA 1248

Search completed: April 3, 2002, 20:58:00
Job time: 42038 sec

us-09-049-696-20.rn

Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:22:08 ; Search time 1321.64 Seconds

(without alignments)
1935.021 Million cell updates/sec

Title: US-09-049-696-20

Perfect score: 2983
Sequence: 1 GAATCAGGAGGAGGTAC.....AAATGTAACCACTGGCTA 2983

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq.1101.*

1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	3111	20	AAZ09840 Human membrane spa
2	2966.2	99.4	3311	22	AAI29502 C902P determined c
3	2828.4	94.8	2854	22	AAH34879 Human colon cancer
4	2828.4	94.8	2854	22	AAH34879 Human colon cancer
5	2814.2	94.3	2867	22	AAH33285 Human secreted pro
6	2813.8	94.3	2825	22	AAH46124 Human C1CA1 CDNA,
7	2793.6	93.7	3109	22	AAH35019 Human colon cancer
8	2743.4	92.0	2745	20	AAH81927 Human ICACC-1 nucl
9	2738.8	91.8	2742	22	AAH46102 Human C1CA1 coding
10	1761.6	59.1	2843	22	AAH46120 Mouse Gob-5 CDNA,
11	1757.6	58.9	2931	20	AAH81925 Murine ICACC-1 nuc

12	1733	58.1	2739	22	AAH46101	Mouse Gob-5 coding
13	1304	43.7	3265	21	AAZ65095	Membrane-bound pro
14	1304	43.7	3265	22	AAH92092	Human PRO1124 CDNA
15	1304	43.7	3265	22	AAH44241	Human PRO1124 (UNO
16	1022.8	34.3	2616	21	AAH64335	Clone 2516888 of a
17	790.8	26.5	878	18	AAH45884	Human colon specific
18	790.8	26.5	878	19	AAH16672	Polynucleotide seq
19	704.4	23.6	1802	21	AAH98067	Human colon cancer
20	704.4	23.6	1802	22	AAH33192	Human colon cancer
21	554.6	18.6	2784	20	AAZ24658	Human lung tumor a
22	554.6	18.6	2784	21	AAH81926	Human lung cancer-
23	552.2	18.5	3951	20	AAH24653	Human lung cancer-
24	552.2	18.5	3951	21	AAH24653	Human lung cancer-
25	552.2	18.5	8031	21	AAH65950	Human lung cancer-
26	531.4	17.8	3156	21	AAH251625	Human membrane cha
27	512.6	17.2	3190	20	AAH81926	Human ICACC-2 nucl
28	433.8	14.5	2454	22	AAH82881	Human tumor-associ
29	366	12.3	3362	21	AAH65896	Human lung cancer-
30	364.4	12.2	3362	22	AAH24657	Human lung tumor a
31	326	10.9	486	22	AAH46121	Mouse Gob-5 CDNA h
32	323.8	10.9	401	21	AAH77755	CDNA encoding huma
33	323.8	10.9	401	22	AAH28493	Colon tumour relat
34	241.2	8.1	255	16	AAH22483	Human gene signatu
35	228.2	7.7	576	21	AAH16018	Human colon cancer
36	221.4	7.4	595	21	AAH16018	Human colon cancer
37	200.8	6.7	618	21	AAH16019	Human colon cancer
38	183.4	6.1	611	21	AAH16022	Human colon cancer
39	168.6	5.7	742	21	AAH16028	Human colon cancer
40	151.6	5.1	375	22	AAH65887	Novel human polynu
41	128.8	4.3	585	20	AAH24629	Human lung tumor a
42	128.8	4.3	590	21	AAH65868	Human lung cancer-
43	124.2	4.2	546	20	AAH24626	Human lung tumor a
44	124.2	4.2	546	21	AAH65865	Human lung cancer-
45	113.6	3.8	936	22	AAH8252	Oligonucleotide D1

ALIGNMENTS

RESULT 1	AAZ09840	standard; CDNA; 3111 BP.
ID	AAZ09840	
AC	AAZ09840	
XX		
DT	26-NOV-1999	(first entry)
DE		Human membrane spanning protein MSP-5 CDNA fragment 2.
XX		
KW		Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;
KW		neoplastic disorder; immunological disorder; reproductive disorder;
KW		MSP-5; ds.
OS	Homo sapiens	
XX		
PN	W09946380-A2	
XX		
PD	16-SEP-1999	
XX		
PF	09-MAR-1999	99WO-US05073.
XX		
PR	13-MAR-1998	98US-0039064.
XX		
FA	(INCY-) INCYTE PHARM INC.	
XX		
PT	Tang YT, Bahdan O, Lal P, Hillman JL, Yue H, Corley NC;	
XX	Geigley KU, Kaser MR, Baughn MR, Shah P;	
XX	WPI: 1999-551409/46.	
DR	F:PSDB: AAY33298.	
XX		
PT	New human membrane spanning proteins used to, e.g. prevent and treat	
PT	neoplastic disorders -	

XX Example 1: Page 80-81; 81pp; English.

XX This invention describes novel human membrane spanning proteins (MSPs),
CC and the polynucleotides encoding them. The products of the invention are
CC used to diagnose, prevent and treat neoplastic, immunological and
CC reproductive disorders. This sequence encodes a human membrane spanning
CC protein MSP-5 fragment.

XX Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match 100.0%; Score 2983; DB 20; Length 3111;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GAATACACAGGAGATGATGACAGCATGGGGCCATTAGATGCTGTCATCTTGATT 60
DB 10 gaatacagaaggagatgacagcaatgggcccatttaagatctctgtctcatcttgat 69
QY 61 CTTCACCTCTAGAAAGGGCCCTGATTAATTCATCTCATCTGACAGACAAATGCTAT 120
DB 70 cttaacctctagaaggccctgagtaattcaatctacagctgagaaacaatgagctat 129
QY 121 GAAGCATTTGCTGTCATGAGACCCCAATGTCAGAAAGATGAACACTCATTTCAACA 180
DB 130 gaagcatctgtctgacatgcaccccaatgagcagaagaatgaaacatctcaacaa 189
QY 181 ATTAAGGACATGATGACCCAGCATCTGTATCTGTTTGAAGCTACAGAAAGCATTT 240
DB 190 ataaagacatgagtgaccacagcatctgtatctgtctgtaagagctaacagaaagatct 249
QY 241 TATTTCAAAAATGTTCCATTTTGTATCTCTGAAAATGGAAGACAAAGCTGATATG 300
DB 250 tattcaaaaatggttcatttgcattctctgtaaacacagaaagacaaagctgacatg 309
QY 301 AGACAAAACCTTGAGACCTACAAAATGCTGATGTTCTGTTGCTGAGCTCTACTCTCA 360
DB 310 agacaaaacctgagacactacaacaaatgctggtctggtctgagctactctctca 369
QY 361 GGTAAATGATGAAACCTTACATGAGCATGGCAACTGTGAGAGAGAGGCTGAAGATG 420
DB 370 ggtaatgataaaccttaacctacagatgagcaactgagagagagaggttaagagatc 429
QY 421 CACCTCACCTCGATTTTCATTTGAGAAAAGTTAGTGATGATGACCAAGATGAG 480
DB 430 caacctcacctcgattctcatgcaagaaagtttagctgaataagacacaaagttag 489
QY 481 GCATTTGTCATGATGAGGCTCATCTACATGGGAGATGATTTGAGAGATGAATATGAT 540
DB 490 gcatttgtcatgagtgagggctcatctacatgagtgaggatattgacgagtaacataatgat 549
QY 541 GAGAAATTTCTATTTCATATGAGAAATACAAAGCTGAAGTGTTCACAGATTTACT 600
DB 550 gagaatcttactatctacaaatggaataacaaagcagtaagatgttcagcaagttact 609
QY 601 GGTAAATGATGAAAGATGTCAGGAGAGAGCTGTTACACAAAAGATGACATTC 660
DB 610 ggtacaaatgataagagatgagtgagagagcagtgtaacacaaagatgacatctc 669
QY 661 AATAAAGTAAAGAGACTGTATGAAAAAGATGACTTTGTTCTCCAAATCCCGCAGAG 720
DB 670 aataaagtaacagactctatagaaaaagatgagttgttcttccaaatcccgccagag 729
QY 721 GAGAAAGCTTTCTATTGTTTGCAACAATGTTGATTTCTATGTTGAATTTCTAGAA 780
DB 730 gagaagctctctataatgttgcacacatgtgattctataatgtgaaatctgtaacaga 789
QY 781 CAATAACCAACAAAGAGCTCAAAACAAAGCAAAATGCAATCTCCGAGACACA 840
DB 790 caataacacacaaagagagctcacaacaaagcaaaatgcaatctccgaagcaca 849
QY 841 TGGAAAGTATCCGTGATTTCTGAGACTTTTAAAGAAACCACTCTTATGACACAGCCA 900
```

```
DB 850 tggaaagtacccgtgattcttgaggaacttaagaaaaaacctctctatgacaaacagcca 909
QY 901 CCAATCCACCTTCTCATTTGCTGACATTTGACAAAGAAATTTGTGTTAGTCTTGAC 960
DB 910 ccaatccacactctctcatcttcagatctgcaagaatgtgtttagctcttgac 969
QY 961 AAATCTGAAAGCATGGGACTGTGAACCGGCTCAATGATGTAATCAAGAGCCAGCTT 1020
DB 970 aaatctgaaagcatgggactgtgaacccgctcaatgactgaatcaagcagccagctt 1029
QY 1021 TTCCCTGACAGACAGTTGAGCTGGGCTCCGTTGGTTGGATGATTTGACAGTGTCT 1080
DB 1030 ttccctgtgacagaaagcttgagctggttccctggttggttgatgacattgcaagtgct 1089
QY 1081 GCCCATGTACAAAGTGAACATCATACAAATTAACAGTGGAGCTGACAGGACACATCGCC 1140
DB 1090 gcccatgtacaagtgaaactatacagataaacagtgagcagtgacagagacaactgccc 1149
QY 1141 AAAGATTTACCTGCAGCAGCTTCAGAGAGGACGTCATCTGACACCGGCTTCGATCGGCA 1200
DB 1150 aaagatltacctgcagacagcttccagagagagagctccatctgcagcggcttgatcgca 1209
QY 1201 TTTACTGTATTTAGGAAAGAAATATCAACTGATGATCTGAATTTGTGCTGACGGAT 1260
DB 1210 tttaactgtatagaaagaaataltcccaactgataatgtgactgaaatgtgtctgtagat 1269
QY 1261 GGGGAGCAACACTATTAAGTGGTGTCTTAAGAGAGTCAACAAAGGAGGATCATATC 1320
DB 1270 ggggagcaaacactataaagtggtgtcttaacagagttcaacaaagtggtgccaatc 1329
QY 1321 CACACAGTGCCTTTGGGCGCCCTCTGACAGCTCAAGACTAGAGAGCTGTCCAAAATGACA 1380
DB 1330 cacacagtcgctttggggccctctgacagctcaagactaagagagagctgtccaaaatgaca 1389
QY 1381 GGAGGTTTACAGACATATGCTTACAGATCAAGTTCAGAAATGCGCTCATTTGATGCTTTT 1440
DB 1390 ggaggtttacagacatagctctcagatcaagttcaagaaatggtcctcatgtgacttct 1449
QY 1441 GGGGCCCTTTGATCAGGAAATGAGAGCTGTCTGACAGCTCCATTCAGCTTGAGAGTAAG 1500
DB 1450 gggggccctttgatcaggaatgagagctgtctcctcagagctccatccagcttgagagtaag 1509
QY 1501 GCATTAACCCCTCAGAACAGCAGCAGTGAATGAGACAGTATGTCGACAGCACCGTG 1560
DB 1510 gcattaacccctcagaacagcagcagtgagatgagacagtgatcgtgacagacaagctg 1569
QY 1561 GGAAGGACACTTTGTTCTTATACCTGACAAACGACGCTCCGCAAAATGCTTCTGCG 1620
DB 1570 ggaaggaacactgttctctataccctcgacaaagcagctccccaatctctctctg 1629
QY 1621 GATCCAGTGGACAGAACCAAGTGCTTTGTATGACAAAACACCAAAATGGCTTAC 1680
DB 1630 gatccagtgagcagaagaagtggtctgtgtagtgagcaaaaacacaaatgacctac 1689
QY 1631 CTCCAATTCACAGGACTTTGCTTAAGTTGGACATTGGAATACAGTCTCAACGAAGCTCA 1740
DB 1640 ctccaatctcagagattgtctaagtttgacacttggaatacagctgcaagcaagctca 1749
QY 1741 CAACCTTGACCTGACTGTACGTCGCTCGGTGCTCAATGTACCTCTCTCAATTACA 1800
DB 1750 caaccttgacccctgactgtcagctgcccgtggttccatgttaacctgtccccaatca 1809
QY 1801 GTGACTTTCAAAAGACAGACAGACACGCAAAATTTCCCAAGCCTCTGTGATGTTATGCA 1860
DB 1810 gtgactttccaaaagacagacagacacagcaaatctcccaagcctctgtgattatgca 1869
QY 1861 AATATTGCGCAAGAGCCCTCCCAATTTCTCAGGGCAGGTGACAGCCCTGATGAATCA 1920
DB 1870 aatattgccaagagagccctcccaatltccagggccagtgtaacagcccttaatgaaaca 1929
QY 1921 GTGAATGCAAAAACAGTGTACTTGAACACTGATTAATGAGACAGTGTGATGCTACT 1980
```


Db 1390 gtgaatggaanaaacagtcacttggaactgataatgagcagtgctgact 1989
 Qy 1981 AAGATGAGGCTGTCTACTCAAGGTATTTTCACACTTATGACGCAATGGTATACAT 2040
 Db 1990 aagatgacgctgtctactcaaggtatctcaactatgatacgaatggtatagcgt 2049
 Qy 2041 GTAAAGTGGGCTCTGGGAGAGACTTAACGACGACGAGAGAGTGTATCCAGCAG 2100
 Db 2050 gtaaaagtcgctgctgaggaatgaaacgacgacgagagtgatccccagcag 2109
 Qy 2101 AGTGGAGCATGTACATCTGGCTGGGATTTGATGATGATGATGATGATGATGATGAT 2160
 Db 2110 agtggagcactgatacaccctgctgctgctgctgctgctgctgctgctgctgct 2169
 Qy 2161 AGACCTGAATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 Db 2170 agactggaatataaagatgataatgataacccaagcagtgctgctgctgctgctgct 2229
 Qy 2221 TCGGAGGCTCATTTTGTGGCTGTGATGTCCCAATGCTCCCATACCTGATCTTCCCA 2280
 Db 2230 tcggaagctcatctgtgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2289
 Qy 2281 CCTGGCCAAATCACCGACCTGAAAGCGGAAATTCACGGGGGAGCTCATATATCTGANT 2340
 Db 2290 cctggccaaatcacccgacctggaagcggaaatlcacgggggagctcatatcatctgact 2349
 Qy 2341 TGAGAGCTCTCTGGGGATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
 Db 2350 tggagagctctctggggatgataatgataacccaagcagtgctgctgctgctgctgctgct 2409
 Qy 2401 AGTACAGTATTTTGTGATCTGAGACAGCAAGTCAATGAACTCTTCAAGTGAATCTACT 2460
 Db 2410 agtaacagatctctgatactcagagacaagtcaatgaaatctctcctcaagtgatact 2469
 Qy 2461 GCTCCATCCCAAGAGGCAACCTGAGGAGGCTTTTGTGATGATGATGATGATGATGATGAT 2520
 Db 2470 gctccatcccaagagagcacaactcggagagcctctctgctgctgctgctgctgctgct 2529
 Qy 2521 ACTTTGAAAATGGCAGATCTTTTTCATTTGCTATTCAGCTGTTGATGATGATGATGAT 2580
 Db 2530 actttgaaaatggcagatctctctcctgctgctgctgctgctgctgctgctgctgctgct 2589
 Qy 2581 AATTCAGAAATATCCAAATTCGACGAGTATCTTTTGTGATGATGATGATGATGATGATGAT 2640
 Db 2590 aaatcagaataatccaaatcagagatctctgctgctgctgctgctgctgctgctgctgct 2649
 Qy 2641 GAGACACCTAGCTGATGAAACGCTGCTCTGCTCTTCAATATTCATATCAACAGCAGC 2700
 Db 2650 gagacacctagctgatacgaacgctgctgctgctgctgctgctgctgctgctgctgctgct 2709
 Qy 2701 ATTCTGGCATTCATCTTTTAAAATTTATGTTGAGAGTGTGATGATGATGATGATGATGAT 2760
 Db 2710 attcctggcatctcaatctttaaataatgagagtgatgagagtgatgagagtgatgagagtgat 2769
 Qy 2761 ATAGCCTTAGGCTGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
 Db 2770 atagccttaggctgataatcttgcagataaataaataatcattccctcttcttcttctg 2829
 Qy 2821 TTAATTAATTTTCTAAATGATTTTATGATCTCTGTTAGGGGCGATTAATTAATGAT 2880
 Db 2830 ttataaattcttcaataatgataatcttgcagatctcgttggggcgatataacaaatgata 2889
 Qy 2881 ATAGTACATTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
 Db 2890 atagatcattataactataatgataatctcgttggggcgatataacaaatgataatcttctg 2949
 Qy 2941 TTCCGTAGGGGCGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2983
 Db 2950 ttccgttagggcgatataaataatgataatctcgttggggcgatataacaaatgata 2992

RESULT 2
 AA129502

Db AA129502 standard; cDNA; 3311 BP.
 Qy AA129502;
 Db 12 OCT-2001 (first entry)
 Qy C902P determined cDNA sequence.
 Db Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 Qy Immunogenic; gene therapy; vaccine; colonic cancer; ss.
 Db Homo sapiens.
 Qy WO200149716-A2.
 Db 12 JUL-2001.
 Qy 29 DEC-2000; 2000WO-US35596.
 Db 30 DEC-1999; 99US-0476296.
 Qy 10 JAN-2000; 2000US-0480321.
 Db 15 FEB-2000; 2000US-0504629.
 Qy 06 MAR-2000; 2000US-0519444.
 Db 19 MAY-2000; 2000US-0575251.
 Qy 25 JUN-2000; 2000US-0609448.
 Db 26 AUG-2000; 2000US-0649811.
 Qy (CORI) CORIXA CORP.
 Db Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;
 Qy King GE, Wang T, Jiang Y;
 Db WPI: 2001-441847/47.
 Qy Colon tumor associated proteins and nucleic acids useful for the
 Db prevention, diagnosis and treatment of colonic cancer -
 Qy Claim 2; Page 425-426; 472pp; English.
 The present invention describes colon tumour associated proteins (I) and
 the polynucleotides (II) that encode them. (I) have cytostatic activity.
 (I) and (II) can be used in gene therapy and vaccine production. (I) and
 (II) may be used in the prevention, diagnosis and treatment of diseases
 associated with inappropriate colon tumour associated protein (TCAP)
 expression, such as colonic cancer. For example, (I) and (II) may be
 used to treat disorders associated with decreased expression by
 relieving mutations or deletions in a patient's genome that affect the
 activity of TCAPs by expressing inactive proteins or to supplement the
 patients own production of them. Additionally, (II) may be used to
 produce the TCAP proteins, by inserting the nucleic acids into a host
 cell culturing the cell to express the protein. (II) and its
 complementary sequences may also be used as DNA probes in diagnostic
 polymerase chain reaction (PCR) and hybridisation assays to detect and
 quantify the presence of similar nucleic acids in samples, and
 therefore which patients may be in need of restorative therapy. (I) may
 also be used as antigens in the production of antibodies against TCAPs
 and in assays to identify modulators of TCAP expression and activity.
 Anti-(I) antibodies and antagonists may also be used to down regulate
 TCAP expression and activity. The anti-(I) antibodies may also be used
 as diagnostic agents for detecting the presence of TCAPs in samples
 (e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
 and AA124494 to AA124523 represent nucleotide and amino acid sequences
 given in the exemplification of the present invention.
 Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;

Query Match 99.4%; Score 2966.2; DB 22; Length 3311;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2979; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 Qy 1 GAAATCACAGGAGATGTACAGCAATGGGGCATTAAAGATTCTGTCTGATCTTGGATT 60

Db 328 ggaatcacaggagatgtagacatgagggccattaaagatctgttcatcttgalt 387
OY 61 CTTCACCTTTAAGAGGGGCGCTGATTAATTCATCTACCTAGCTGAACAACATGGCTAT 120
Db 388 cttaacctcttagaagggcccggaatctcaatctcaatgcttagacaacaatggtctat 447
OY 121 GAAGGCAATTCGCTTGGCAATGAGCCCAATGTCACAGACATGAACACTTCATTCACAA 180
Db 448 gaagagcttgctgtgtaacgcccccaatctgtccagaagaatgaacacatcttaacaa 507
OY 181 ATAAGAGACATGCTGACCCAGCATCTCTGATCTGTTAGCTTACAGAAAGCATTT 240
Db 508 ataaagacatgtagaccagacccctgtagcttctgagctcaagaagaagcgatctc 567
OY 241 TATTTCAAAATGTTGGCATTTGATTCCTGAAACATGGAAGCAAGGCTGATGTG 300
Db 568 talttcaaaaatgttgcaatcttgatctcctgaaacatggaagaagaagctgactatgtg 627
OY 301 AGACCAAACTTGAGACCTTACAAAATGCTGATTTCTGCTGCTGAGTCTACTCCCA 360
Db 628 agaccaaaacttgagacctacaaaatgctgaigtctgtgtgtagcttaactctcca 687
OY 361 GGTATGATGAACTTACCTAGAGAGATGGCAACTGTGGAGAGAGAGGTGAAGGATC 420
Db 688 ggtaatgtagaccctacacatgagcagatggtgcaactgtgagagagaggtgaaagatc 747
OY 421 CACCTCACTCTGATTTGATTCAGAGAAAAGTTAGCTGAATTTGACCAACAGGTAG 480
Db 748 caccctacatctgattatctgcaagaaaagttagctgaatctgagcccaagagtag 807
OY 481 GCATTTGTCATGATGGGCTCATCTATAGATGGGAGATTTGAGAGTACATATATGAT 540
Db 808 gcaatctgccaagagtggtccatctacagatgagagagatcttgacagataatgat 867
OY 541 GAGAAATTTCTACTTATCCATGGAAGAAATACAGCATGATGTTGACGAGATTTACT 600
Db 868 ggaataatctactatctcaatgaaatatacagcagtagaatgtctcagcaggtatct 927
OY 601 GGTACAAATGTTGTAACAAAGTGTAGGAGGAGGAGCTGTTACCAAAAAGATGACATTC 660
Db 928 ggtacaaaatgtagtaagaagatgctcaagagagagcgtgtacacaaaagatgacatctc 987
OY 661 AATTAAGTAAACAGACTTATGAAAAAGATGTGATTTGTTCTCAATCCCGCAGACG 720
Db 988 aataaagttaacgagctctatgaaaaaagatgtgagttgtctccaaatcccgcaagag 1047
OY 721 GAGAGGCTTGTATTAATGTTTGCACAACTGTTGATTCATAGTGAATTCGTACAGAA 780
Db 1048 gagaagctctctataatgtttgcacacatgttgatctcatagttgaatctcgtaagaa 1107
OY 781 CAATAACCAACAAAGACCTCCAAACAAAGCAAAATCAAAAATGCAATCTCGAAGCACA 840
Db 1108 caaaaccaacaagaagaagctccaacaagaacaaatcgaatctccgaagcaca 1167
OY 841 TGGGAAGTGAATCCGTGATTTGAGACTTTAAGAAAACACATCTATGACACAGCCA 900
Db 1168 ttggaagtgatcgtagtcttgagagactttaaagaaaacacatctctgaacaacaagcca 1227
OY 901 CCAAAATCCCACTTTCATGCTGAGATGTGACAAAGATTTGTTAGTCTTGAAC 960
Db 1228 cbaaatcccaactctcaatctgtcgcagatitggaacaagaatgtgtgttagtccctgac 1287
OY 961 AATATGGAAGCATGGGAGCTGTAAACCGCTCAATGACATGATTAAGCAGCCAGCTT 1020
Db 1288 aatcttggaagcatgagcagctgttaaccgcctcaatcgatgaatlaaacagagccagctt 1347
OY 1021 TTCTCTGTGAGACACTGTGAGCTGGGGTCTGGTGGATGGTGAATTTGACATGTGCT 1080
Db 1348 ttctctgtcagagacgttgagctgagggctcctggtctgagagcttgacattgacagtctc 1407
OY 1081 GCCCATGTACAAAGTGAACATATACAGATAAAGTGGCAGTGAAGGAGCAGACTCGCC 1140
Db 1408 gcccaatgataaagttagaactcatcagataaaacagtgagcagtgagacagacactcgcc 1467
OY 1141 AAAAATTAACCTGACGACGCTTTCAGAGGAGCGTCATCTGACAGGGGCTTCATCGGCA 1200
Db 1468 aaaaatctacccgacagcgtctcaaggaagagctccacatctgcagcggctctgctggca 1527
OY 1201 TTATGATGATGAGAAAGAAATATATCAACTGATGATGATGATGATGATGATGAT 1260
Db 1528 ttactgtgattaggaagaataatctcaactgtatgtgactgaaatgtgtctgtgagagat 1587
OY 1231 GGGGAAGACACACTATAAGTGGTGTAAAGAGGTCCTTAAAGAGGTCATCTC 1320
Db 1588 ggggaagacaacactataagtggtgtcttaacgaggtcaacaagaatggtgcatctc 1647
OY 1321 CACAGCTGCTTTGGGGGCTCTCTCAGCTCAAGAACTAAGAGAGCTGTCCAAAATGACA 1380
Db 1648 caacagctcgcttggtggccctctgagctcaagactaagagagctgtccaaatgaca 1707
OY 1331 GGAGGTTTACAGACATATGCTTACATCAAGTTCAGAACATGCTCATTTGATCTTTT 1440
Db 1708 gtaggtttacagacataatgtctcaagatctcaagatctcaagaacatggtccatgtatctt 1767
OY 1421 GGGGCGCTTTCATCAGAAATGAGGCTGTCTCAGCGCTCCATCCAGCTTGAAGTAAG 1500
Db 1768 ggggccccctcacaggaatgtagctgtctctcagcgtctcatccatccatgtgagtag 1827
OY 1501 GGATTAACCTCCAGAAACAGCCAGTGAATGACACAGTATGTGTGACAGACCGTGTG 1560
Db 1828 ggaatccctccagaaacagcagctgagatgagacagtgatctgtgacagcccggtg 1887
OY 1561 GGAAGGACACTTGTGTTTATCACCCTGAGACACACGCTCCCAAAATCTTCTGTG 1620
Db 1898 ggaagagacatctgtctctcaacatgagacagcagcctcccaaatctctctctg 1947
OY 1621 GATCCAGTGGAGAGAAAGCAAGGTGCTTTGATGAGCAAAAACCAAAAATGGCTTAC 1680
Db 1948 gatccagtgagaaagcaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2007
OY 1681 CTCCAAATCCAGGCAATGCTTAAGCTTGGACATTTGAAATACAGTGTGCAAGCAACTCA 1740
Db 2008 ctccaatcccaagcattgtctaaagttgagcttggaataaagctgtgcaagaagctca 2067
OY 1741 CAACCTTGAACCCGACTGTACGCTCCCGTGCCTCAATGCTCAATGCTCAATTTAC 1800
Db 2068 caaaccttgaccctgactgtgacgtctccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2127
OY 1801 GTGACTTCCAAAAGCAACAGAGACACCAAGCAATTTCCCGCCCTCGTAGTTATGCA 1860
Db 2128 gtgacttccaaaagcaacaggaacacagcaaatctcccgccctctgtgtgtgtgtgt 2187
OY 1861 AATTTGCCCAAGAGGCTCCCAATTTCTAGGCGCAGTGTCAAGCCCTGATTTGAATCA 1920
Db 2198 aatatctgcaagaagagctccccaatctccagggcaggtgtcaagggccctgtatgtca 2247
OY 1921 GTGAATGGAAGAAACAGTTTACCTTGAACACTGATGATTAATGAGAGAGTGTGATGCTACT 1980
Db 2248 gtgaatgaaagaagatcttgaactgtgaactgtgaatgaatgaatgaatgaatgaatga 2307
OY 1991 AAGATGAGCGGTGCTCTCTCAAGGATTTTCAACTTATGACAGATGATGATGATGAT 2040
Db 2308 aagatgagcgtgtgtctcaactcaaggtatctcaactcaatgaacagatgtgtgtgtgtgt 2367
OY 2041 GTAAAGTGGCGGCTCTTGGAGAGATTAACGACACGACGAGAGATGATACCCACAG 2100
Db 2368 gtaaaagtggcggtctgtggagaggttlaacgacgacgaggaagatgtatcccaagcag 2427
OY 2101 ACTGAGACACTGTATATACCTGCTGATTTGAGATGATGATGATGATGATGATGATGAT 2160
Db 2428 agttagagactgtatataccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2487
OY 2161 AGACCTGAAATTAATTAAGATGATGTTCAACAGCAAGAGTGTGTTTCAAGAGAAATCC 2220
Db 2488 agactgaaatataaagaatgtatgttcaaaccaagcaaggtgtgtgtgtgtgtgtgtgt 2547

OY	2221	TCGGGAGGCGCATTTTGGCTTCTGTGAATGTCGCCAAATAGCTCCATATCGGATCTGTGCCA	2280
Db	2348	TCGGGAGGCGCATTTTGGCTTCTGTGAATGTCGCCAAATAGCTCCATATCGGATCTGTGCCA	2607
OY	2281	CCCTGGCCAAATTCACCGCAGTCGAAAGGGGGAAATTCACGGGGGACGTCCTATTATCTGACT	2340
Db	2608	CCCTGGCCAAATTCACCGCAGTCGAAAGGGGGAAATTCACGGGGGACGTCCTATTATCTGACT	2667
OY	2341	TGGACAGCTCCCTGGGGATGATTATGACCAATGGAAACAGCTCACAGATPATATTCGAATA	2400
Db	2668	TGGACAGCTCCCTGGGGATGATTATGACCAATGGAAACAGCTCACAGATPATATTCGAATA	2727
OY	2401	AGTACAGTATTTCTGTGATCTCAGACAGCAAGTTCATGAAATCTCTTCAGTGAATATCTACT	2460
Db	2728	AGTACAGTATTTCTGTGATCTCAGACAGCAAGTTCATGAAATCTCTTCAGTGAATATCTACT	2787
OY	2461	GCTGTATCCCAAGGAAGCAACTCTGAGGAAGTCTTTTGTATTAACCGAAACAACTT	2520
Db	2788	GCTGTATCCCAAGGAAGCAACTCTGAGGAAGTCTTTTGTATTAACCGAAACAACTT	2847
OY	2521	ACTTTTGAAAATGGCAGACAGATCTTTTCATTTGCTATTCAGGCGTGTGATTAAGTGCATCTG	2580
Db	2848	ACTTTTGAAAATGGCAGACAGATCTTTTCATTTGCTATTCAGGCGTGTGATTAAGTGCATCTG	2907
OY	2581	AAATTCACAAATATTCACAATTCGACAGCAGATATCTTTGTTTATTTCTTCACACAGCTCCGCA	2640
Db	2908	AAATTCACAAATATTCACAATTCGACAGCAGATATCTTTGTTTATTTCTTCACACAGCTCCGCA	2967
OY	2641	GAGACACCTTGTGCTGTGATGAAACGTGTGCTCTTGCTTAATATTCATATCAATCAAGCACC	2700
Db	2968	GAGACACCTTGTGCTGTGATGAAACGTGTGCTCTTGCTTAATATTCATATCAATCAAGCACC	3027
OY	2701	ATTTCCTGGCATTCACATTTTAAAAATTTATGTGGAAGTGGATAGGAAGAACTGCAGCTGTCA	2760
Db	3028	ATTTCCTGGCATTCACATTTTAAAAATTTATGTGGAAGTGGATAGGAAGAACTGCAGCTGTCA	3087
OY	2761	ATACCTTAGGGCGCGAATTTTGTGCACATTAATAATAATCAATTCATCTTTTGTGGA	2820
Db	3088	ATACCTTAGGGCGCGAATTTTGTGCACATTAATAATAATCAATTCATCTTTTGTGGA	3146
OY	2821	TTTATAAATTTTCTTAAAAATGTATTATTAAGACTTCCTGTAGGGGGGATPATCAATGTAT	2880
Db	3147	TTTATAAATTTTCTTAAAAATGTATTATTAAGACTTCCTGTAGGGGGGATPATCAATGTAT	3206
OY	2881	ATATGACATTTTATCTTAATATGATTTCTGTAGGGGGCGATATCTAAATGATTTTATGAC	2940
Db	3207	ATATGACATTTTATCTTAATATGATTTCTGTAGGGGGCGATATCTAAATGATTTTATGAC	3266
OY	2941	TTTCTGTAGGGGGCGATAAAAATTAATTCGTAACCACTGGCTGA	2983
Db	3267	TTTCTGTAGGGGGCGATAAAAATTAATTCGTAACCACTGGCTGA	3309
RESULT: 3			
AAH34879 standard; cDNA; 2854 BP.			
AAH34879:			
03-SEP-2001 (first entry)			
Human colon cancer antigen encoding cDNA SEQ ID NO:1961.			
Human: colon cancer; colon cancer antigen; diagnosis: detection;			
colorectal carcinoma; chromosome 1; ss.			
Homo sapiens.			
M0200122920-A2.			
05-APR-2001.			
28-SEP-2000, 2000MO-US26524.			

XX	29 SEP-1999;	99US-0157137.
PR	03 NOV-1999;	99US-0165280.
XX	(HUMA-) HUMAN GENOME SCI INC.	
XX	Ruben SM, Batash SC, Birse CE, Rosen CA;	
EL	WPL; 2001-235357/24.	
FR	P-PSDB; AAG75474.	
FT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,	
PT	useful for preventing, diagnosing and/or treating colorectal cancers -	
XX	Claim 1; Page 3462-3463; 9803pp; English.	
FS	AAG72943 to AAH37195 and AAG73514 to AAG77788 represent human colon	
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where	
CC	the proteins are collectively known as colon cancer antigens. The colon	
CC	cancer antigens have cytostatic activity and can be used in gene	
CC	therapy and vaccine production. N and P may be used in the prevention,	
CC	diagnosis and treatment of diseases associated with inappropriate P	
CC	expression. For example, N and P may be used to treat disorders	
CC	associated with decreased expression by rectifying mutations or deletions	
CC	in a patient's genome that affect the activity of P by expressing	
CC	inactive proteins or to supplement the patients own production of P.	
CC	Additionally, N may be used to produce the colon cancer-associated ps,	
CC	by inserting the nucleic acids into a host cell and culturing the cell	
CC	to express the proteins. N and P can be used in the prevention, diagnosis	
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204	
CC	and AAH77789 represent sequences used in the exemplification of the	
CC	present invention.	
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were	
CC	missing at time of publication, meaning no sequences are present for	
GC	SEQ ID NO:1027 to 1052, 7921 and 7922.	
SQ	Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;	
Query Match	94.8%; Score 2828.4; DB 22; Length 2854;	
Best Local Similarity	99.8%; Pred No. 0;	
Matches 2832; Conservative	0; Mismatches 6; Indels 0; Gaps 0	
Gy	1 GAAATCACAGGAGATGTACAGCAATGGGGCCATTAAAGATTCTGTTCATTTGGATT 60	
Dd	11 gaatcacacgggagatgttacagaacaatgggccatttaaggatctcgttcatcttgat 70	
Gy	61 CTTCACCTTCTTGAAAGGGGCCCTGACTAATTCACATCATTCAGCTGAACAACAATGCGTAT 120	
Dd	71 cttcaccttctaagaaggccctgtagtaattcaatcactcagtgaacaacaatggtctat 130	
Gy	121 GAAGGATGTGTGGTCATTCGACCACCCANMGCCGAAGAATGAATAACATCANTCAACA 180	
Dd	131 gaaggcatgttgttgcatacgaccaccaatgycagaagaatgaaacacactcatacaaa 190	
Gy	181 ATTAAGGACATGGTAGCCAGGCATCTCTGTATCTTTGAAGCTCAGAGAAAGCCATT 240	
Dd	191 ataaggacatggtgacccaagcatctcgtatctgttgaagctacaggaagcgat 250	
Gy	241 TAATTTCAAATAATGTTGCCATTTTTGATTCCTGAACAATGGAACAAAAGCTGATATG 300	
Dd	251 taattcaanaatgctgcattttgatcttcctgaacaatggaagaagaagctgactatgt 310	
Gy	301 AGACCAAACTGTAGACCTACCAAAAAATGTGATGTCCTGGTTCGAGTCTACTCTCCA 360	
Dd	311 agaccaaaacttgagaccatacaaaaaatgctgagttcgttgcagttctactcccca 370	
Gy	351 GGTAATGATGAACCCCTACACTGAGCAGATGGGCACTGTGAGAGAAAGGTTGAAGATC 420	
Dd	371 ggtaatgtagaacctcactacactgagcaatggtggaactgtggagagaagggtgaaggatc 430	
Gy	421 CACCTCACTCTGATTTCAATTCAGGAAAAAAGTTAGCTGAATATGACCACCAAGGTAG 480	

Db 431 caccctcctgattctatctgacaggaaaaagtagctgaatatacgaccacaagtag 490
QY 481 GCATTTGTCATAGTGGGCTCATTCACATGGGAGATATTGACGAGTACATTAATGAT 540
Db 491 gcatcttgcacagtaggctcctcctcaagatgaggatatttgacagatcaaatatgat 550
QY 541 GAGAAATTTCTACTTATCCAAATGGAAGAAATCAAGAGTAAGTTCACAGATATTACT 600
Db 551 gagaatctactatcccaatggaagatacaagcagtaagatgctcagcaggtatctact 610
QY 601 GGTACAAATGTAGTAAGAGTGTCAAGGAGGAGCGTGTACACCAAAATGACATTC 660
Db 611 ggtacaaatgtagtaaaagtagcaggaggcagctgtacaccaaagatgacatc 670
QY 661 AATAAATGACAGGACCTATGAAAAAGATGTGAGTTGTTTCCAAATCCGCGACAG 720
Db 671 aataaagtacagagctcctatgaaaaagtagtgattgtctcccaatccgcagag 730
QY 721 GAGAGGCTTCTATATGTTTGCACAAATGTTGATCTATGTTGTAATCTGTACGAA 780
Db 731 gagaagcttctataatgattgtcacaacatgltgattctatgtagtcaatctcgtacagaa 790
QY 781 CAAACCCACAAAGAAAGTCTCAACAGCAAAATCAAAATGCAATCTCCGAGACCA 840
Db 791 caaaaccacaacaaagaaagctccaaacaaagcaaaaaatgcaatcccgagacaa 850
QY 841 TGGGAAGTATCCGTGATTTGTGAGACTTTAAGAAAACCTCTATGACACACACCA 900
Db 851 tgggaagtagatccgtgattctcgtgagactttaagaaaacaaatccatctgacacacagca 910
QY 901 CCAAAATCCCACTTCTCTGTTGCTGAGATGAGCAAAATGTTGTTGTTGTTGTTGAC 960
Db 911 ccaaatcccaactctctcctgctgagatgagcaaaatgtagtctgttctgtctgtgac 970
QY 961 AAATGTGAGAGATGGCACTGTGTAACCGCTCAATCGACTGAATCAAGACGCGCAGTT 1020
Db 971 aaatctggaaagatgagcagactgtagaacgcctcaatcgactgaatcaagcagcgact 1030
QY 1021 TTTCTGCTGACAGAGTTGAGCTGGGCTCTGCTGGTGGATGGTGAATTTGACAGTGTCT 1080
Db 1031 ttcctgctgacagacagcttgagctgaggctgagctgagctgagctgagctgagctgagct 1090
QY 1081 GCCCATGTGCAAAATGTAATCTATACAGATTAACAGTGGCAGTGCAGAGGACACACTGCGC 1140
Db 1091 gcccatgtacaagtagaactcctacagaaagtagcagtagcagtagcagtagcagtagcag 1150
QY 1141 AAAAGATTACTGTGACAGAGCTTTCAGAGGAGAGCTCATCTGACAGGCGCTTCGATCGGCA 1200
Db 1151 aaaagattactgtcagcagctcctcagaaagtagcagtagcagtagcagtagcagtagcag 1210
QY 1201 TTTTACTGTGATTAAGAAATATCCAACTGATGATGTAATTTGCTGCTGACGAT 1260
Db 1211 ttttactgtatctagaagaataatacctcaactgtagctgaaatgtgtgtgtgtgtgtgtgt 1270
QY 1261 GGGGAAGACACATTAATGAGGCTTTAAAGAGTCAAAAGTGTGCTCATATC 1320
Db 1271 ggggaagacaacaactaagtagtggtctttaaagagtagcacaagtagtggtcctcactc 1330
QY 1321 CACACAGTGTGTTGGGCGCTTCGAGCTCAAGAACTGAGAGAGTGTGCAAAATGACA 1380
Db 1331 cacacagtgctgttgaggcctctgagctcagaaacttagaagagtagtgcacaaatgaca 1390
QY 1381 GAGAGTTTACAGACATATGCTTACATCAAGTTACAGAACATGAGCTCATTTGCTTTT 1440
Db 1391 ggaagtttacaagacatagtctcagatcaagttcagaaacatgagctcattgtgtctt 1450
QY 1441 GGGGCGCTTTCATCAGGAAATGAGCTGTCTCAGCGCTCATCAAGCTTGAGATTAAG 1500
Db 1451 ggggcgcttctcatcagaaatgagctgtctcagcgctcctcagcttgagtagaag 1510
QY 1501 GGAATTAACCTTCAGAACACAGAGTGTGATGATGAGCAAGTGTGACACACCTG 1560
Db 1511 ggaataaccttcagacagcagtagtagaagtagcacagtagtctgtgacacagcacctg 1570

QY 1561 GGAAGAGACCTTTGTTCTTATACCTGAGCAACGACGCTCCCAAAATCTTCTG 1620
Db 1571 ggaagagacacttgattctctatcaccgagcaagcagctccccaatctctctg 1630
QY 1621 GATCCAGTGGACACAGACAGAGTGGCTTTGTAGTGACAAAAACCAAAATGGCTTAC 1680
Db 1631 gatccagtaggacagaaagtagtgcttgtagtggacaaaacacaaaatgctctac 1690
QY 1681 GTCCAAATCCAGCATTTGCTTAAGTTGGACCTTGGAAATACAGTCTCCAGCAAGCTCA 1740
Db 1691 cccaatccagcagcttgtagtgaagcttgacatggaataagctcagcaagcagca 1750
QY 1741 CAAACTTGACCTGACTGTACGCTCCGCTGCTCAATGTCTACCTGCTCAATTACA 1800
Db 1751 caaacttgaccttgagctgacgctccgctgctcagctgctcagctcctcctcaat 1810
QY 1801 GTGACTTCCAAAACGAAACAGACACACAGCAAAATTCACGCTCTGTGATTTATGCA 1860
Db 1811 gtgacttccaaaacgaaacagacacacagcaaaatccccaagccttgtagttagtaca 1870
QY 1861 AATATGCGCAAGGAGCGCTCCCAATTCAGAGCCAGTGCACAGCCCTGATTAATCA 1920
Db 1871 aatattgcgaagagcctccccaatccagggcagtgtaacagccttgatgacatca 1930
QY 1921 GTGAATGAAAAACAGTTACTTGAATCTGATTAATGAGACAGTCTGATGCTACT 1980
Db 1931 gtgaatgaaaaacagttactccttgagactcagtagaatagagcagtgctgtagtact 1990
QY 1981 AAGGATGACGGTGTCTACTCAAGTATTTCAACAACTTATGACAGAAATGTGATCACT 2040
Db 1991 aaggatgacgggtgtctactcaagtagtattcaacatctgaacagaaatgtagatacag 2050
QY 2041 GTAAAGTGGGCGCTGAGAGAGTAAAGCAGCCAGACGAGAGTGTATCCAGCAG 2100
Db 2051 gtaaaagtggcgctctgaggagagtagtaacgacagcagagagtagtaccacag 2110
QY 2101 AGTGGACACTGTACATACCTGCTGAGTTGAGAAATGATTAATTAATGAAATCCACCA 2160
Db 2111 agtggacactgtacatacctgctgtagtgaatgaatgaatgaatgaatgaatgaatgaatga 2170
QY 2161 AGACCTGAATTTAATAGATGATGTCAACACAGCAAGCAATGTGTTTACGACAGACATCC 2220
Db 2171 agacctgaatataaagtagatgtagtcaacacacagaaagtagtctcagcagaacatcc 2230
QY 2221 TCGGAGAGCTCATTTGTGCTGTGATGTCCCAATGCTCCATACCTGATCTCTCCCA 2280
Db 2231 tcggagagctcattgtggtcttgtagtcccaatgtctccatctgactccttccca 2290
QY 2281 CCTGGCCAAATCACCGACCTGAAAGGGGAAATTACAGGGGAGCTCATTAATGATGACT 2340
Db 2291 cctggccaaatcacagcctgaaagcgaaatctcaaggggagctcctcaatctgact 2350
QY 2341 TGGACAGTCTCTGGGAGTATGACCAATGAAACAGCTCAACAAGTATATTCGATA 2400
Db 2351 tggacagtctctgggagtagatctagacacatgaaacagctcctcaatctgactcctcaat 2410
QY 2401 AGTACAAATTTCTTATCTCAGAGACAACTTCAATGAATCTCTCAAGTAAATCACT 2460
Db 2411 agtaaaatcttctgactcagagacaaagtcaatgaatctctcaagtagaataact 2470
QY 2461 GCTTCATCCCAAGGAGACCACTGAGGAATCTTTTGTGTTAAACGAAACAT 2520
Db 2471 gcttca-cccaagaaagcaaatctcagaaagtagtcttctgttaaaccgaaacatc 2530
QY 2521 ACTTTGAAATGACAGATCTTTTCAATGCTATTCAGCTGTGATTAAGTGTGATCTG 2580
Db 2531 actttgaaatgacagatcttctcattctgactcagctgtgataagtagcagctg 2590
QY 2581 AATACGAATATTCACAAATTCACAGAGTATCTTGTATTCACCAACAGCTCCGCA 2640
Db 2591 aatcagaataatcccaatctgacagtagtattctgttattctcccaagacacgcgca 2650

|||||
Db 851 tgggaagatgacgctgagctcctgagagccttaagaaccaccctcctatgacacacagcca 910
OY CCAATTCACACCTCTCTCTGCTGAGATTTGGACAAAGATTTGTTGTTGCTCTTAC 960
Db 911 ccaatccacccctccctcctgctgagatctgacaagaatctgctgcttccctcctgac 970
OY 961 AAATCTGGAGCATGGCAGCTGTAAACCGCTCAATCGACTGAAATCAAGAGCCAGCTT 1020
Db 971 aaatctggaagcatgagcctgtaaccgctcctcaatcgactgaatcaagcagcgactt 1030
OY 1021 TTCTCTGCTGAGACAGTTGAGCTGGGCTCTGGGTTGGGATGGTGACATTTGACAGTCT 1080
Db 1031 ttcctgctgcagacagctgagctgaggctcctgagctgagctgagcttgcagctgct 1090
OY 1081 GCCCATGTACAAAGTGAAGCTATACAGATTAACAGTGGCAGACAGGACACATCCGC 1140
Db 1091 gcccatgtacaagaagtgacatcatalaagaataacagtgagcagtgacaaggaacacacgc 1150
OY 1141 AAAAGATTACTGACAGCAGCTTTCAGAGAGGAGCAGTCCATCTCGAGCGGCTTCGATCGCA 1200
Db 1151 aaagattactctgcagcagctctcagagagagcgtccatctgcagcagcgcttcgcagca 1210
OY 1201 TTTACTGTGATTAGAGAAATATCCAACTGATGATGTGAATTTGCTGCTGACGAT 1260
Db 1211 ttctactgtagttaggaagaaatcccaactgtagtctgaaatctgctgtagcagat 1270
OY 1261 GGGGAGAGCAGCAGTATTAAGTGGGCTGTTTAAAGGTCAAAAGAGTGCCATCATC 1320
Db 1271 ggggaagagacacatataagtgagtgcttcaagagtgcaaaagaagtgagcctcctc 1330
OY 1321 CACACAGTGGCTTTGGGGCCCTGACAGCTCAGAACTAGAGAGCTCTGCCAAATGACA 1380
Db 1331 cacacagcgtcttgaggcctctgcagctcagaactagaggagctgctccaaaatgca 1390
OY 1381 GGAGGTTTACAGACATNGCTTCAGATCAATTTGAAACATGGCTTATGATGATCTTT 1440
Db 1391 ggaagttcaagacatagctcagatcaagltcagaacaatgagcctcatctgactctt 1450
OY 1441 GGGGCCCTTTCATCAGAAATGAGAGCTGTCTCTCAGCGCTCATCAGCTTGAGAGTAAG 1500
Db 1451 ggggccccttcatcagagaaatgagcgtgctcctcagcgtccatccatcagctgagagtaag 1510
OY 1501 GGATTAACCTTCAGAACAGCAGCAGTGGATGAATGGCAGATGCTGGAGACAGCGTG 1560
Db 1511 ggaataaccctccagaagaacagctgagatgagcacagtgatgtagaagcagcgtg 1570
OY 1561 GGAAGAGCAGCTTTGTTCTTATCACCTGGAACAGCGAGCTCCCAAAATGCTCTGG 1620
Db 1571 ggaagagacacttgcttcttaltacccggaacagcagcctcccaaaatccctcctc 1630
OY 1621 GATCCAGTGGACAGAAAGCAGAGTGGCTTTGTAGTGACAAAACACCAAAATGGCTTAC 1680
Db 1631 gatccagtgagcagaagaagcagtgcttgtagtgagcaaaaacacaaatgagcctac 1690
OY 1681 CTCCAAATCCAGCAGCATTTGTAAGTTGGCACTTGGAATAACAGTCTGCAAGCAAGTCA 1740
Db 1691 ctccaaatcccaagcagcatgctaaagtgtagcacttgaaatacagctgcaagcaagctca 1750
OY 1741 CAAACCTTGACCTGAGCTGTACGTCGCTGCGTCCGATGCTACCCCTCCCAATTCAC 1800
Db 1751 caaaccttgaccctgagctgtagcgtccggtgctccaatgctacccctcccaattca 1810
OY 1801 GTGACTTCCAAAGAACAGAACAGACACAGCAAAATTCACAGCCCTCTGTAGTTATGCA 1860
Db 1811 gtgacttccaaaagaagaagaacacagcaaatccccaagcctctgtagttagttagca 1870
OY 1861 AATATTGCCAAGAGAGCTCCCAATTTCTCAGGGCCAGTGTACAGCCCTGATGAATCA 1920
Db 1871 aatatctgcgaagagcctcccaatctcagggccaggtgcacagccctgatgaatca 1930
OY 1921 GTGATGGAAAAAGTACTGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
|||||

|||||
Db 1931 gtagtggaaaaacagcttaactctggaactactgataatgagcagctgctgtagtact 1990
OY 1981 AAGATGACGGGTCTACTACAGGATTTTTCACACTTTATGACAGAAATGATGATGATGAT 2040
Db 1991 aagatgacgggtctactaccagagatcttcaaacactatgaaacagatgagtagaatacag 2050
OY 2041 GTAAAAGTCCGGGCTCTGGAGAGACTTAACGACAGCAGACAGAGAGATNACCCAGCAG 2100
Db 2051 gtaaaagtcgggctctgaggagtagtaacgacagcagcagcagagtagatccccaagag 2110
OY 2101 AGTGACAGCAGTGTACATACCTGCTGCTGATTTGAGATGATGAATTCATTCGATCCACA 2160
Db 2111 agtggagactgtataatccctgctgtagatgagaatgataatcaatggaatccacaca 2170
OY 2161 AGACCTGAATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db 2171 agacctgaatataaagagagtagtctcaacaacagcagcagcagcagcagcagcagcagc 2230
OY 2231 TCGGAGGCTCATTTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2280
Db 2241 tcgggagcctaatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 2290
OY 2281 CCTGCCAAATCACCAGCAGCTGAGCGGAAATTCACGGGGCAGTCTCATTAATCTGACT 2340
Db 2291 cctggcccaatccacagcagcctgaaagcggaaatccacggggcagctcctaatctgact 2350
OY 2341 TGCACAGCTCCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db 2351 tggacagctcctgaggagtagtaagcagagaaagcagcagcagcagcagcagcagcagc 2410
OY 2411 AGTACAGATTTCTGATCTCAGAGACAGTTCATGATCTCTTCAAGTGAATGATGATGAT 2460
Db 2421 agtacaagtatcttgatctcagagagcaagttcaatgatactctcctcaagtgataactact 2470
OY 2461 GCTCTACCCCAAGAGAGCAACCTCTGAGAGAGCTTTTGTGTTAAACGAGAAACAT 2520
Db 2471 gctccaccccaagagagcagcagcctgaggaagctcttctgtttaaaccagaaacat 2530
OY 2521 ACTTTTGAATAATGGCAGCAGATCTTTTCAATGCTATTCAGCGGTGATGAAGTGCATG 2580
Db 2531 aacttgaataatgagacagatcttctcactgctcagcgtctgtagtaaggtcagatctg 2590
OY 2581 AAATCAGAAATATCCACATTTGACAGAGTATCTTTGTTATTCCTGACAGACTCGGCA 2640
Db 2591 aaatcagaataatccaaatgacagagtagtcttcttattccctccacagactccgca 2650
OY 2641 GAGACACCTAGTCCGATGAAAGCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700
Db 2651 gagaacactagctcagagaaagcgtcgtcctcgtcctcctcaataltacatacaacagc 2710
OY 2701 ATTCTGACATTCACATTTTAAATTTATGTAAGTGAATGATGATGATGATGATGATGATGAT 2760
Db 2711 attcctgcatcattacattttaaaataltatggaagtagatagaggaactcagctgca 2770
OY 2761 ATAGCTTAGGCTGATTTTGTGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2820
Db 2771 atagcctagggcgtgaatttctgtagtaataataataatcattcattccttcttga 2830
OY 2821 TTATTAATTTTCTTAATA 2838
Db 2831 ttataaaaaaaaaa 2848
|||||

RESULT 5
AAH33285
ID: AAH33285 standard; cDNA: 2867 BP.
XX. AAH33285;
AC. (first entry)
XX. C3-SEP-2001 (first entry)
XX. Human colon cancer antigen encoding cDNA SEQ ID NO:341.
XX

KM Human: colon cancer: colon cancer antigen: diagnosis: detection;
KM colorectal carcinoma: chromosome 1; ss.
OS Homo sapiens.
XX
XX MO200122920-A2.
XX
XX 05-APR-2001.
PD
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI
XX WPI: 2001-235357/24.
DR P-PSDB; AAG73854.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX
XX Claim 1: Page 2452-2453; 9803PP: English.
PS
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37195 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
CC
XX
XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;
SQ

Query Match 94.3%; Score 2814.2; DB 22; Length 2867;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2830; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GAAATCACAGGAGATGTACAGCAATGGGGCCATTAAAGAGTCTGTGTCATCTGATT 60
DB 14 gaattacacagggagatgacagcaatggggccattaaagatctgtgtacatcttgatt 73
QY 61 CTTCACCTTCTAGAAGGGCCCTGAGTAATTCATCTACATTCAGCTGAACAACATGGCTAT 120
DB 74 cttaacctctgaagggccctgagtaattcctcatcagctgagcaacaatgacctat 133
QY 121 GAAGCATTTGCTTGCAATTCGACCCCAATGTGCGAGAGATGAACACTCATTTCAACA 180
DB 134 gaagcatctgtcttgcaattcgaccccaatgtgcgagagatgaacactcatcaaca 193
QY 181 AATAAGCATATGATGACCAAGCATCTGTATCTGTTGAGACTACAGAAAGCATTT 240
DB 194 aataagcatatgatacgaagcatctgtatctgtttgagactacagaaagcatatt 253
QY 241 TATTTCAAAAATGTTGCTATTTGATCTCTGAACATGGAAGCAAAAGGCTGACTATGTC 300
DB 254 tatttcaaaaatgttgctattttgatctctgaacatggaagcaaaaggctgactatg 313

AGACCAAACTTGAGACCTACCAAAAATGCTGATGTTGCTGCTGAGTACTCTCA 360
1314 agacaaaacttgagacctacaaaatgctgattgctgtgtgcttactctctca 373
QY 1 GGTATGATGATACCCCTACATGAGCAGATGGCGAAGCTGTGAGAGAAAGGATG 420
DB 1314 ggtatgatatgataccctacatgagcagatggcgagctgtgagagaaaggatg 433
QY 421 CACCTGACCTCTGATTTGATTCAGAGAAAAAGTTAGCTGAATATGAGACCAAGTAGG 480
DB 434 cacttgacctctgatttcatcttgagaaaaagttagctgaatatgagaccaaagtagg 493
QY 481 GCATTTGCTCAGTGGGGCTCATCTACGATGGGAGATATTTGACGATCAATATGAT 540
DB 494 gcatttgctcagtggggctcatctacgagtgaggagatttgacgagtaataatgat 553
QY 541 GACAATTCATCTTATTCATGAGAGAAATTCAGACACTAAGATGTTGACGATTTACT 600
DB 554 gagaattctactatccatcgaagaaatacgaacagtaagatgttcaagagttact 613
QY 601 GTTACAATGATGATTAAGAAGTGTGAGGGAGGAGGAGCTGTACACCAAAAGATGACATTC 660
DB 614 gttaaatgatatgattaaagatgtgagggagggagagctgtacaccaaaagatgacattc 673
QY 661 AATAAGTAAACAGGACTCTATGAAAAAGATGAGTTGTTGCTCCAAATCCCGCAGAGC 720
DB 674 aataagtaaacaggactctatgaaaaagatgagttgttcttccaaatccgcagagc 733
QY 721 GACAGGCTTCTATATGTTTGCACAACATGTTGATTTCTATGTTGATTTCTGATGAA 780
DB 734 gacaggcttctatattgttgcacaacatgttgatttctatgttgatattctgtacgaa 793
QY 781 CAATACCAACAAAGAGGCTCAACCAACCAATCAAAATGCAATCCCAAGACACA 840
DB 794 caataccaacaagagagctccaaacaacaacaacaatgcaatcccaagacaca 853
QY 841 TGGAGATGATCCGCTGATTTGAGACTTTAAGAAAACCTCTATGACAAACAGCCA 900
DB 854 tggagatgatccgctgatttgagactttaagaaaacctctatgacaaacagcca 913
QY 901 CCAATTCGCCACTTCTCATTTGCTGACAGATTGCAAAAGATTGTGTTTATGCTTGAC 960
DB 914 ccaatttcgccacttctcatTTGCTGACAGATTGCAAAAGATTGTGTTTATGCTTGAC 973
QY 961 AATTCGAGAGCAATGGGAGCTGTGAACCGGCTCAATGATGATCAAGACAGCGACTT 1020
DB 974 aatctggaagcattgagcagctgtgaacccggctcaatgattgatacaagagcagctt 1033
QY 1021 TTCCTGCTGACAGACTGAGCTGGGGTCTGGGTTGGATGGTGAATTTGACAGTGT 1080
DB 1034 ttcctgctgacagactgagctggggctggggttggaatggatggaatttgctgagcag 1093
QY 1031 GCCCATGTAACAAGTACATCATACAGATAACAGT- GCGAGTGACAGAGACACACTCGC 1139
DB 1034 gcccatgtacaagtacatcatatcagataaacagt-gcgagtgacagagacacactcgc 1153
QY 1140 CAAAGATTTACCTGACAGACTTCAGAGAGGAGGAGCTCATGTCAGCGGCTTGATGAGCG 1199
DB 1134 caaagatttacctgacagacttcagagagggagagctcatgtcagcggttgatgagcgc 1213
QY 1200 ATTTACTGTGATTAAGAAATATCCAACTGATGATGATGAAATTTGCTGCTGACGGA 1259
DB 1214 atttactgtgattaagaaatattccaaactgattgattgctgctgagcagga 1273
QY 1260 TGGGGAAGACACATCTAATAGTGGTGCTTTAACGAGGTAAACAAAGTGTGCTCATCAT 1319
DB 1274 tggggaagacacatctaatagtggtgcttttaacgaggtaaacaaagtgtgctcatcat 1333
QY 1320 CCACACAGTGGTTTGGGGCTCTGACAGTCAAGAACTGAGAGAGTGTCCAAATGAC 1379
DB 1314 ccacacagtgtgttggggctctgacagtcaagaaactgagagagtggtccaaatgac 1393
QY 1380 AGGAGGTTTACAGACATATGCTTCAGATCAAGTTCAAGAACATGGGCTCATGTGCTTT 1439

D	1394	agggaggttacagacatactgcttcagatcaagtctcgaatacgaatggccatcttgatgcttt	1453
Q	1440	TGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCGCTCCATCAGCTTGAGAGTAA	1499
D	1454	tggggcccttcattcagaaatgagctctctctccagctccatccagcttgaagtaa	1513
Q	1500	GGGATTAAACCTCCAGAAACGCGCAGTGGATGGAATGGCACAGTGAATCGTGGACAGCAGCT	1559
D	1514	gggattaaacctccagaaacagccagtgagatggatgcagatgacatcgatgaagaacgct	1573
Q	1560	GGGAAAGACACTTGTGTTCTTATCACCTGGACACGACGCTCCCAATCTCTCTG	1619
D	1574	gggaagagacacttgyttcttcttaacacctgagacaagcagctccccaatctctctg	1633
Q	1620	GGATCCCAAGTGGACAAAGCAAGAGTGGCTTGTGATGGACAAAAACCAAAATGGCTTA	1679
D	1634	ggatcccaagtgagaaagaaagaaggtgcttcttgatgagcaaaaacacccaatggccta	1693
Q	1680	CCTCCAAATCCAGAGCAATTCCTTAAGTGTGGCACTTGGAAATACAGTCTGCAGCAAGCTC	1739
D	1694	cttccaatcccaaggaatctgctaagtttgacatcttgaaatacagctctgcaagaagctc	1755
Q	1740	ACAAACCTTGACCTGACTGTCACTGACGTCGCGTGCTCAATGCTACCTGCTCCAAATTC	1799
D	1754	acaaacctgacctgactgtactcactgcctcgctgccaatgtacacctgcccctaattac	1813
Q	1800	AGTACATCCAAAGCAAGCAAGACACAGCAAAATCCCAAGCCGCTGGTGTATATGC	1859
D	1814	agtactctccaaaacgaaacaagaacacccagcaatccccagcctctgtaagttaagc	1873
Q	1860	AAATATTTCGCCAAGAGAGCTCCCAATTTCTCAGGGCCAGTGTACAGCCCTGATTTGAATC	1919
D	1874	aaatatttcgcaagagagcccccacatctcagagcgagtgccaagcccttgatgaatc	1933
Q	1920	AGTAAATGAAAAACAGTTTACCTTGGAACTCGATATGGAATGGAGAGGTGTGATGTGCAC	1979
D	1934	agttaaagaaaaaacagcttaccctctggaactacacggttaatggacagcgatgcagctaac	1993
Q	1980	TAAAGATGACGGTGTCTACTCAAGGATTTTCACACCTTATGACAGGAATGATATACG	2039
D	1994	taagatgacggtgtctactccaagtatttccaacaattatgacaagaatgtagtatacag	2053
Q	2040	TGTAAACTGCGGGCTGTGGGAGAGATTTAAACGACGCAACGGAAGATATACCCACACA	2099
D	2054	tgtaaactgcgggctctggagagagatlaacgaagcacaagagatgataccccaca	2113
Q	2100	GAGTGAAGCACTGTACATACCTGGCTGGATGGATGATGAATGAATACATGGAATCACCC	2159
D	2114	gagtgagacactgtaactactcgtgcgtatgtggaatgtgnaatacaatgtaataccac	2173
Q	2160	AAGACTGGAATTTAATAAGGATGATGTTCAACACAAAGCAAGTGTGTTTACGACAAATC	2219
D	2174	aagactggaatttaataaagatgattgttcaacaacagcaagtggttlaagcaagaatac	2233
Q	2220	CTCGGAGGGCTAATTTGTGGCTCTGATGTGCCAAATGTCCTCCATACCGATCTTGCC	2279
D	2234	ctcggagggctcaatctgtggtctctgtagtgcaccaatgtcccataaccgatactctccc	2293
Q	2280	ACCTGGCCAATTCACGACCTTGAAGGCGGAAATTCACGGGGCAGTCTCATTAATCTGAC	2339
D	2294	acctggccaataccagcactggaagcggaattcacaggggcagctctcataatctgac	2353
Q	2340	TTTGACACGCTCTGGGGATATATATATGACCATGGAACAGCTCACAGTATATATCTGAAAT	2399
D	2354	tttgacacgctctcggggatgattatgacaaatggaaacagctcaaatatataatctgaaat	2413
Q	2400	AAGTACAAAGATTTCTTGATTTTCAGACACAAGTTCAATGATGTCTTCAAGTGAATATACAC	2459
D	2414	aagtacaagatattcttgatcttcagagacaagttcaatgtaatctcttcaagtgaaatacac	2473
Q	2460	TGCTCTCATCCCAAGAGACCAACTGTGAGCAAGTCTTTTGTTTAAACGAGAAACAT	2519

Dd	2474	tgcctcgtccccaagaagcaacgcaactcggaagctcttcttgtttaaacaggaaacat	2533
Cy	2570	TACTTTTGAATAATGGCAGCATCTTTTCATTGCTAATTCAGCGCTGTATTAAGTGCAATCT	2579
Dd	2574	tactttgaaaaattggcacagaatcctttcatatgcatlcaagcigtgtataagtcgatct	2593
Cy	2580	GAAATCACAAAATATTCACATTTGCACGAGATCTTTGTTTATTCCTCCACAGACTCCGCC	2639
Dd	2584	gaaatcagaaaataaccaacatltgcagagatcctttgttatcttccccaagaetccgcc	2653
Cy	2670	AGAGACACCTTAGTCCTCGATGTAAGAAGCTGCTGCTCTTGTCTTAATATTCATATCAACAGCAC	2699
Dd	2674	agagcacacctagctcgtatgnaacgctgcctccttgtccataatatcatacaacagcac	2713
Cy	2770	CATTCCCTGGCACTTCACATTTTAAAAAATATATGTGGAAGGATAGAGAACACGACGCTGTC	2759
Dd	2774	cattcccgtgcatlcaacatltaaaataatlaagtgaagtagaagaacagcagctgctc	2773
Cy	2770	AATAGCCCTAGGCGCTGAATTTTTTTCACAGATATATTAATTAATCATCTCTTTTTTTTG	2819
Dd	2774	aatagccctagggcgtgaatcttttgcagagtaataataataatcatcatcctcttttttg	2833
Cy	2820	ATTATTAATTTTCTTAAAA 2838	
Dd	2834	attataaaaaaaaaaaaaa 2852	
RESULT 6			
AAH46124	FD	AAH46124 standard; cDNA; 2825 BP.	
XX	AA846124;		
AC			
DT	11-SEP-2001 (first entry)		
XX			
DE	Human CLCA1 CDNA, SEQ ID NO:26.		
XX			
KX	Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;		
KW	expression inhibition; antisense therapy; gene therapy;		
RW	chronic obstructive pulmonary disease; bronchial asthma; antilasthmatic;		
LX	ss.		
FM			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	23..2767	
FT		/*tag_a	
FT		/product= "Human CLCA1"	
FT		/trans_except= (pos:476..478, aa:Lys)	
XX			
PX	WO200138530-A1.		
XX			
FD	31-MAY-2001.		
FE	22-NOV-2000; 2000WO-JP08232.		
XX			
PR	24-NOV-1999; 99JP-0333479.		
PR	27-APR-2000; 2000JP-0127589.		
XX			
EA	(TAKE) TAKEDA CHEM IND LTD.		
XX			
PI	Nakanishi A., Morita S;		
ER	NPJ. 2001-355935/37.		
DR	P-PSDB; AAB73716.		
XX			
PT	New antisense nucleotide, useful for treatment and prevention of		
PT	bronchial asthma and chronic obstructive pulmonary disease -		
XX			
PS	Example 5; Page 92-94; 104pp; Japanese.		
XX			
CC	This invention relates to an antisense nucleotide targeted to the mouse		
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart.		

the CLCA1 gene (coding sequence shown in AHA46102). The invention also relates to an antibody specific for the Gob-5 protein, medical and diagnostic compositions containing the antisense nucleotide or the antibody, and methods and kits for screening for compounds which inhibit the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells. The antisense oligonucleotides and antibody are therefore useful for the treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease. The present sequence represents human CLCA1 cDNA.

Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;

Query Match 94.3%; Score 2813.8; DB 22; Length 2825;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2818; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 AATCCAGGAGATGTACGCAATGGGCGCATTTAGAGTTCGTCATCTGATTC 62
DB 1 aatcacaaggagatgtacgcaatgggcccattagaagatctgttcaatcttgatct 60
QY 63 TCACCTTTAGAGGGGCCCTGAGTAATTCATCTACCTGACGTAACAGATGGCTATGA 122
DB 61 tcacctctagaaggggcccttgatattcatctatctgcgtgaacaacaatgctatga 120
QY 123 AGCGATGTGCTGTCATGCAATGCCAATGTGCCAGAGATGAACACTGATTCACAAAT 182
DB 121 agcgatgtgctgtgcatcgaccaccaatgtgccagaagatgaacactatcaacaat 180
QY 183 AAGGACATGAGGACCCAGCATCTCTGATCTGTTGAAGCTACAGGAAAGGATTTA 242
DB 181 aagaagacatgagaccagcatctctglatcgtltagagctacaggaaagatllta 240
QY 243 TTTCAAAAATGTGCAATTTGATTCCTGAACATGCAAGACAAAGCTGACATGTGAG 302
DB 241 ttcaaaaatgtgcatatttgatcttccctgaacaatgaaagaaagcgcacatgtgag 300
QY 303 ACCAAATCTTGAGACCTACAAAAATGCTGATGTTGCTGAGCTACCTCCAGG 362
DB 301 accaaacttgagacctacaaaatgtgatgtctgtgtgctgagctacatctccagg 360
QY 363 TAATGATGAACCTACACTGAGCAGATGGGCAACTGTGAGAGAGGATGAAGATTC 422
DB 361 taatgatgaacctacactgagcagagatgggcaactgtgagagaaagagatccaa 420
QY 423 CCTCATCTGATTTCAATTCAGAGAAAAGTTAGCTGATATGACCAAGATGGGC 482
DB 421 cctcatctctgattcatctatgcaagaaaagttagctgaatatgagccaaagtgag 480
QY 483 ATTGTGCATGAGTGGGCTCATCTACGATGGGAGTATTTGACGAGTACAAATGATGA 542
DB 481 attgtgcatagtggtgctcatctacgatagtgagattgagaggtacataatgata 540
QY 543 GAAATTCATCTATCCAAATGAAGAAATACAGCAGTAAGTATGTCAGAGTATTC 602
DB 541 gaatttcattctatccaaatggaagatacaagaaagatglttcagcaggtatctcg 600
QY 603 TACAAATGATGAAGAAATGTCAGAGGAGGAGCTGTACACCAAAAGATGCATTCAA 662
DB 601 tacaatgtagtaagaagatgtcagagagagcgtgttacccaagaatgacatctcaa 660
QY 663 TAAAGTACAGAGCTCTATGAAGAAAGATGAGTGTGTTCTCAATCCCGCCAGAGGA 722
DB 661 taagttaacagactctatgaagaaagatgtgtgttcttccaatcccgccagcgga 720
QY 723 GAAGCTTCTATATGTTTGGACAATGTTGATCTATAGTTGATGTTCTGTACAAACA 782
DB 721 gaagcttctataatgttggacaacaatgtgactctatagttgaaatctgtacaga 780
QY 783 AAACCAACAACAAGAACTCCAAACAGCAAAATCAAAATGCAATCTCCGAGACATG 842
DB 781 aaaccaacaacaagaagctccaaacaagaataaataatgcaatctccgaagacatg 840
QY 843 GGAAGTGAATCGTGATTTGAGGACTTTAAGAAACCACTCCTATGACAAACAGCCACC 902

DB 841 ggaagtgatccgctgattctgagacttaagaaacacatccatctgacaacacagccacc 900
QY 903 AATTCACACTTTCATTCGTCGAGATTTGACAAAGAAATGTGTGTTTACTCTTGACA 962
DB 901 aatcccaactctcatctgctgacatgagatggaacaagaatgtgtttaaactctgaca 960
QY 963 ATCTGAGAGCATGGCAGCTGTGTAAACCGCTCAATTCGACTGAATCAAGCAGCCAGCTTT 1022
DB 961 atctggaagatgtagctgtgataccgcctcacaatctgactgaatcaagcagcgacttt 1020
QY 1023 CCTGCTGACAGACGTTGAGCTGGGCTCTGGTGGATGGTATGATTAACAGTGTGCTG 1082
DB 1021 cctgctgacagacgcttgagctgggctctggctgtgtgagatgtgacattgacagtgctgc 1080
QY 1083 CCATGTACAAAGTCACTCTACAGATTAACAAACATGTGGCAGAGACAGGACACTCCGCCAA 1142
DB 1081 ccatgtacaagaatgaaactctacagaaacagtggaatgacaaggaacactcgcaca 1140
QY 1143 AAGATTACCTGCAGACGCTTCAGAGGAGACGTCCATCTGCAGCGGCTTCGATCGCAT 1202
DB 1141 aagattacctgcagacgcttcagagaggaagctcatctcgaagcggtcttgatcgcgcat 1200
QY 1203 TACTGTGATTAGAGAAATATCCAACTGATGATCTGAATTTGCTGCTGACGATGG 1262
DB 1201 tactgtgattagaagaataatccaaactgagatcgtgaatgtgtcgtgcgagatg 1260
QY 1263 GGAAGACAAACACTTAATAGTGGGCTTTTACGAGGTCAAAAGTGGTCCATCATCCA 1322
DB 1261 ggaagacaacacttaaatagtggtgtttaaagagttcaacaagaatggtgacatccaa 1320
QY 1323 CACAGTCCCTTTGGGGCCCTCTCAGCTCAAGAACTAGAGAGAGCTGCCAAATGACAG 1382
DB 1321 cacagtcccttggggccctctcagctcaagaaactagagagagctgccaaatgacag 1380
QY 1383 AGGTTTACAGCAATATGCTTCAGATCAAGTTCAAAATGACCTCATTTGATGCTTTGG 1442
DB 1381 aggtttacagcaatattgcttcagatcaaatgctcagaaatgacctatgtactttgg 1440
QY 1443 GGGCTTTTCATCAAGGAAATGAGACTGTCCTGACGCTCCATCAGCTGAGATGAAGG 1502
DB 1441 gggcttttcatcaggaatgagactgtctcagctccatccatcagcttgaagataagg 1500
QY 1503 ATTAACTCTCAGACAGCAGAGTGAATGACACAGTATGATGATGAGACAGCAGCTGGG 1562
DB 1501 attaaactctcagacagagcagtgatgaatgacacagatgctgagacagcagctgg 1560
QY 1563 AAGGACACTTTTGTCTTATATCACTGAGACAGCAGCTCCCAAAATCCTTCTGGGGA 1622
DB 1561 aaggacactttgttcttatacactgagacaagcagctcccaaatcctctctggga 1620
QY 1623 TCCCATGAGACAGAACAGAGTGGCTTTGATGAGGACAAAAACCAAAATGGCCTTACCT 1682
DB 1621 tcccatgagacagaaacagagtggtctttagtggacaacaaaacacaaatggtccact 1680
QY 1683 CCAAAATCCAGAGGATGCTAAGGTGGCACTTGGAAATACAGTCTGCAAGCAGCTCACA 1742
DB 1681 ccaaaatccagagatgctaaagttggtgaccttgaataacagctcgtcaagcaagctcaca 1740
QY 1743 AACCTTGACCTTCAGCTGTCACGTCCCGTGGTCAATGCTACCTGCTTCAATTAACGT 1802
DB 1741 aaccttgaccttcagctgacgtcccggtgctcaaatgtacacctgtcctcaattacagt 1800
QY 1803 GACTTCCAAAGCAACAAGACACGCAAAATTCGCCAGCCCTCTGTGATTTATGCAAA 1862
DB 1801 gacttccaaaagcaacaagacacgcaaaatccccaagccctctgtgattatgcaaa 1860
QY 1863 TATTCGCAAGGAGCCTCCCAATTCAGAGGCGCAGTGTACAGCCTGATTAATCACT 1922
DB 1861 tatctgcgaaggagcctccccaattctcagggcagtgatgacagcctgtatgaatcagt 1920
QY 1923 GAATGAAAAACAGTTACCTTGGAATCTGATGATTAATGAGCAGGTGCTGATGATCTAA 1982

480 GCATTGTCATGAGTGGGCTCATCTACGATGGGAGATTTGACAGTACATAATGA 539
|||||
343 ggcatcttgccatgagtgaggccatcactacgattgaggatattgacataatga 402
|||||
540 TCAGAAATCTCTACTTATCCAAATGGAAGATACAGAGTAAGTGTCCAGCAGCTATTAC 539
|||||
403 tgaagaattctactatccaaatgaagaataaagaagatgcttcagcagattac 462
|||||
600 TGTCAAAATGTAGTAAAGATGTCAGAGAGCAGCTGTACACCAAAAGATGACATTT 659
|||||
463 tggtaacaaatgataagaatgctcaggaggcagctgttaacccaagaatgacacatt 532
|||||
660 CAATTAAGTAAACAGACGCTCTATGAAAAAGAGTGAATTTGTTCTCAATCCCGCAGAC 719
|||||
523 caataaagtacagagactatgaaagaatgtagtctgtctcccaatcccgccagac 582
|||||
720 GGAGAAAGCTTCTATTAATGTTTGCACAAATGATGATGTTGTTGATTTCTGATACAG 779
|||||
583 ggaagaagcttcataatgattgacaaacatgltgattctatgattgaaatctgtacaga 642
|||||
780 ACAAACCCACAACAAAGAGCTCCAACAAAGCAAAATGCAATCTCCGAGACAC 839
|||||
643 acaaaaccacaacaaagaagctccacaacaaatgcaaaatgcaatctccgaagcac 702
|||||
840 ATGGAGATGATCGGTGATTTGAGGACTTTTAAAGAACCACTCTATGACAAACAGCC 899
|||||
703 atggagatgatacgtgattctgaggacttaagaacacactctcatgacaacagacc 762
|||||
900 ACCAAATCCACCTTCTCATGTCGAGATGAGACAAAGATGTCGTTTACTCTTGA 959
|||||
763 accaaatccacactctcactgctcagatlggaacaaatgtagtltgattcattga 822
|||||
960 CAATCTGGAAGACATGGCAGCTGTTAACCGCTCAATCCAGTGAATCAAGCAGCCAGCT 1019
|||||
823 caaatctggaagcatgagcagctgataacccgctccaatcgatgaatcaagcagccagct 882
|||||
1020 TTTCTGCTGACAGACATTTGAGCTGGGGTCTGGGTTGGAGTGGATTTGACAGTGC 1079
|||||
883 ttctccgctgcagacagctgagctgagctgagctgagctgagctgagctgagctgagctg 942
|||||
1080 TGGCCATGTACAAAGTGAATCTATACAGATTAACAGTGAAGTGAAGGAGCAGCTGCG 1139
|||||
943 tgcctcatgatacaagatgaactcaatacaagatgaagatgaagatgaagatgaagatgaag 1002
|||||
1140 CAAAGATTTACCTGACAGCTTTCAGAGAGGAGCTTCATCTGACCGGCTTCGATCGCG 1199
|||||
1003 caaagattacactgagcagcttcagaggagcttcacatctgagaggcttcgacgcgc 1062
|||||
1200 ATTTACTGTGATTAGGAATAATATCAACTGATGATCTGAAATTTGCTGCTGACGGA 1259
|||||
1063 attactctgattaggaagaataatcaactgattgaaatgctgctgacgga 1122
|||||
1260 TGGGGAAGACACACTATTAAGTGGGTGCTTTAAGGAGTCAAAAGTGGTCCACATCAT 1319
|||||
1123 tggggaaagacaacataaagtgagtgcttcaagagtgcaacaaagtgagtgacatcat 1182
|||||
1320 CCACACAGTCTGCTTTGGGGCCCTCTGACAGCTCAAGAACTAGAGAGAGCTGTCAAAATGAC 1379
|||||
1183 ccacacagctgcttggggccctctgacagctcaagaaagagagagctgtccaaatgac 1242
|||||
1380 AGGAGGTTTACAGACATATGCTTCAGATCAAGTTCAGAAATGCTGCTCATTTGATCTTT 1439
|||||
1243 aggaagtttacaagacataagcttcagatcaagatlcagaaacatgagctcatgattctt 1302
|||||
1440 TGGGGCCCTTTATCATGAGAAATGAGACTGCTCTCAGCGCTCATCAGCTTGAAGTGA 1499
|||||
1303 tggggcccttcatcagaaatgagagtgctctcagcgctccatccagcttgaagtaa 1362
|||||
1500 GGGATTAACCTTCAGAAAGCCAGTGAATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1559
|||||
1363 gggattaaacctccagaaacagcagtgagtgaaatggcaagatgtagtggagcagcagct 1422
|||||
1560 GGGAAAGACACTTTGTTTCTTATCATCTGACACGACAGCTCCCAAAATCTTCTG 1619
|||||

1423 gggaaaggacacttgcttcttcaacctggaacagcagccctcccaaatctctctg 1482
|||||
1630 GGAATCCAGTGGACAGACAGAGTGGCTTTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1679
|||||
1433 ggaatcccaatggaacagaagaaggtagcttctgtagtgaagaaacacaaatggccta 1542
|||||
1680 CCTCCAAATCCAGGAGTGGCTTAAGTGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1739
|||||
1513 cctccaaatcccaagcagcttcgaagtgtagcttgcattggaataaagcttcgaagcagctc 1602
|||||
1740 ACAACCTTGACCTGACACTGTACGCTCCGCTGCTCAATGCTACCTGCTCAATTTAC 1799
|||||
1603 acaacctgacactgactgactgacgctccgctgagctccaaatgactcccaatgact 1662
|||||
1800 AGTACCTTCCAAAGCAGACAGACACGACCAATCCCAAGCTCTGCTGATTTATGC 1859
|||||
1653 agtgaacttcccaaaacagaacagacacagcaaatcccaagcctctgtagtattgac 1722
|||||
1860 AATATTTGCGCAAGAGAGCTCCCAATTTCTCAGGCGCAGTGCACAGCCCTGATTTGATC 1919
|||||
1783 aatatctgcgaagagagccctcccaatctcagagcgagtgctacagccctgattgac 1782
|||||
1920 AGTGAATGGAATAACAGTTACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1979
|||||
1733 agtgaatggaataaacagcttacccttggaactgagtaatgagcagtgctgattgac 1842
|||||
1980 TAAGATGACGCTGCTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2039
|||||
1893 taagatgagcagtgctacacaaagatcttcaacacttaagacgaatgagtagataag 1902
|||||
2040 TGTAAAGTGGCGGCTCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2099
|||||
1973 tgtaaagtgagggccctcagagagtgatgaacgaacgaagcagagagtgatgaacgaacga 1962
|||||
2100 GAGTGAAGCAGTGTACATACCTGCTGATTTGAATGATGAATGATGAATGATGAATGATGAAT 2159
|||||
1953 gagtgaagcagctgataactctgctgagatgagaaatgagaaatgagaaatgagaaatgagaa 2022
|||||
2160 AAGACCTGAATTAATTAAGATGATGTTCAACCAAGCAAGTGTGTTGACGAGAAATC 2219
|||||
2023 aagaactgaatlaataaagaaatgattcaacaaacagcagagtgatctccagacatc 2082
|||||
2240 CTGGAAGGCTCATTTTGAGCTTCTGATGTCCCAATGCTCCCATACCTGATCTCTTCC 2279
|||||
2093 ctggagaggtcatcttgctgcttcgattgctccaaatgctcccatcactctctcc 2142
|||||
2280 ACCTGGCCCAATATACGAGCTGGAAGGCGGAATTCACGGGGGAGTCTATTATCTGAC 2339
|||||
2143 acctggcccaatatacagacactggaagcagaaatlcacgggggagcttcaatactgac 2202
|||||
2340 TTGACAGCTCTCGGGGATGATTAATGACCAATGGAACAGCTCAAGATATTCATTTGGAAT 2399
|||||
2403 tggagacgctcctggggatgattatgacccaatggaacagcacaagatataatctcgaat 2262
|||||
2400 AAGTAAAGTATTTCTGATCTCAGAGACAAATGATGATGATGATGATGATGATGATGATGATGAT 2459
|||||
2253 aagtaagaatcttcgattctcagagacaaagctcaatgaaatctcctcaagtgaaatactac 2222
|||||
2460 TGTCTCATCTCCAAAGAGGACCAACTCTGAGGAAGTCTTTTGTTTTAAACAGAAAAAT 2519
|||||
2323 tgtctcatctccaaagagacacacttgaggaagcttcttctttaaaccagaacaaat 2382
|||||
2520 TACTTTGAAATGGACAGATCTTTTCAATGCTATTCAGCGCTGATTAAGTGCATCT 2579
|||||
2403 tactttgaaatggaacagacatcttctcatctcattcagagctgtgataaagtgatct 2442
|||||
2580 GAATCAGAAATATCCAAACATTTGCAGAGTATTTGTTTATTTCTCCACAGACTCGCC 2639
|||||
2443 gaatacagaataatccaaatctgacagagatcttcttatactcccaagcctccgctc 2502
|||||
2640 AGAGACACTGATCTGATGAAGACGCTGCTGCTTGTGCTTAATTCATATCAACAGAC 2699
|||||

805 AACAAACAAATCAAAAATGCATCTCCAGACACATGGAGATGATCCGTGATCTTGAG 864
|||||
781 aacaaacaaatcaaaaaatgcaatctccgaagcacatggaaatgatacgttcttgag 840
|||||
865 GACTTAAAGAAACCACTCTTATGACACAGCCCAATCCCACTTCTCATTTGCGG 924
|||||
841 gacttaagaaacacccctctatgacaacacagcccaaatccaccctctcatctgctg 900
|||||
925 GAGATTGACAAAGAAATTTGTTTATGCTTGAACAAATCTGGAAGATGAGGACTGTG 984
|||||
901 cagattggacaaagaatttggttttaagtccttgacaatctggaagcatgagactggt 960
|||||
985 AACCCGCTCAATGACTGATGATCAAGCAGCCAGCTTTCTCTGCTGACAGACATTTGAGCTG 1044
|||||
961 aacccgctcaatcgactgatacaagcagccagcttctctgctgagacagttgagctg 1020
|||||
1045 GGGTCCTGGTGGATGGTGTGACATTTGACAGTGGCCCACTGTAAGTGAACCTCA 1104
|||||
1021 gggctcctgggttggatggatgacatctgacagtgctgccaatgtaaaagtgaactcata 1080
|||||
1105 CAGATTAACAGTGGCAGTACAGAGGACACATCGCCAAAAGATTACTGACAGAGCTTTC 1164
|||||
1081 cagataaaacagtgagcagtgacagggacacacccgcaaaagattactgacagacttca 1140
|||||
1165 GAGAGGACGTCCATCTGACGCGGCTTGATCGGCAATTTACTGTGATTAAGAAATAT 1224
|||||
1141 ggaaggagcgtccatctgcagcggtctcgatcgcatctactgattagaaagaatat 1200
|||||
1225 CCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1284
|||||
1201 ccaactgataatgataatgataatgataatgataatgataatgataatgataatgata 1260
|||||
1285 TGCCTTAACAGAGTCAACAAAGTGGTGCATCATCATCATCATCATCATCATCATCAT 1344
|||||
1261 tgccttaaacagtgataacaaagtggtgcatacatcaacacagtcgcttggggccctc 1320
|||||
1345 GCAGCTCAAGAACTAGAGAGAGCTGTCCAAAATGACAGAGGTTTACAGACATATGCTTCA 1404
|||||
1321 gcaagctcaagaactagaagagctgttcaaaatgacagaggttcaagacatagcttca 1380
|||||
1405 GATCAAGTTGACAAACAAATGGCTCATGATGATGATGATGATGATGATGATGATGAT 1464
|||||
1381 gatacagttcgaacaaatgagcctcatctgcttggggcccttcatcaagaataatgta 1440
|||||
1465 GGTGTCTCTCAGCGCTCCATCCAGCTTGAAGATTAAGATTAACCTCCAGAACACCCAG 1524
|||||
1441 gctgtctctcagcgctccatccagcttgagatgaaggttaacctccagaacacccag 1500
|||||
1525 TGGATGAATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1584
|||||
1501 tggatgaatgacagatgataatgataatgataatgataatgataatgataatgataatg 1560
|||||
1585 AACCTGGAACAAAGCAGCTTCCCAAAATCCCTTCTGGAATCCCAAGTGAACAAAGAGT 1644
|||||
1561 accctggacaaacagcagcctcccaaatccctcttgagatcccaatgagaaagcaaggt 1620
|||||
1645 GGCCTTGTAGTGAACAAACAAACAAATGACCTACTACCTCAAAATCCAGCATTTGATAG 1704
|||||
1621 ggccttgttagtgagacaaacacacaaatgacctactcccaatcccaagcatgtgtaag 1680
|||||
1705 GTTGGCACTTGGAAATACGTCTGACAGCAAGCTCAAAACCTTGACCTGACTGTACG 1764
|||||
1681 gtggcacttggaataatcagctgcaagcaagctcacaacacctgacactgactgcaag 1740
|||||
1765 TCCCGGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1824
|||||
1741 tcccggtgctcaaatgatacagctccctcccaatcagatgactcccaacagcaagagac 1800
|||||
1825 ACCAGCAAAATTTCCAGCCCTGTGATGATTTATGCAAAATTTCCGCAAGAGAGCTCCCA 1884
|||||
1801 accagcaaatccagccctctgtgattatgcaaatatctgccaagagagctcccca 1860
|||||
1885 ATTTCAGGCGCAGTGTCAACAGCCCTGATTTGATGATGATGATGATGATGATGATGAT 1944
|||||

1885 ATTTCAGGCGCAGTGTCAACAGCCCTGATTTGATGATGATGATGATGATGATGATGAT 1944
|||||
1801 accagcaaatccagccctctgtgattatgcaaatatctgccaagagagctcccca 1860
|||||
1825 ACCAGCAAAATTTCCAGCCCTGTGATGATTTATGCAAAATTTCCGCAAGAGAGCTCCCA 1884
|||||
1741 tcccggtgctcaaatgatacagctccctcccaatcagatgactcccaacagcaagagac 1800
|||||
1765 TCCCGGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1824
|||||
1681 gtggcacttggaataatcagctgcaagcaagctcacaacacctgacactgactgcaag 1740
|||||
1705 GTTGGCACTTGGAAATACGTCTGACAGCAAGCTCAAAACCTTGACCTGACTGTACG 1764
|||||
1621 ggccttgttagtgagacaaacacacaaatgacctactcccaatcccaagcatgtgtaag 1680
|||||
1645 GGCCTTGTAGTGAACAAACAAACAAATGACCTACTACCTCAAAATCCAGCATTTGATAG 1704
|||||
1561 accctggacaaacagcagcctcccaaatccctcttgagatcccaatgagaaagcaaggt 1620
|||||
1585 AACCTGGAACAAAGCAGCTTCCCAAAATCCCTTCTGGAATCCCAAGTGAACAAAGAGT 1644
|||||
1501 tggatgaatgacagatgataatgataatgataatgataatgataatgataatgataatg 1560
|||||
1525 TGGATGAATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1584
|||||
1441 gctgtctctcagcgctccatccagcttgagatgaaggttaacctccagaacacccag 1500
|||||
1465 GGTGTCTCTCAGCGCTCCATCCAGCTTGAAGATTAAGATTAACCTCCAGAACACCCAG 1524
|||||
1381 gatacagttcgaacaaatgagcctcatctgcttggggcccttcatcaagaataatgta 1440
|||||
1405 GATCAAGTTGACAAACAAATGGCTCATGATGATGATGATGATGATGATGATGATGAT 1464
|||||
1321 gcaagctcaagaactagaagagctgttcaaaatgacagaggttcaagacatagcttca 1380
|||||
1285 TGCCTTAACAGAGTCAACAAAGTGGTGCATCATCATCATCATCATCATCATCATCAT 1344
|||||
1261 tgccttaaacagtgataacaaagtggtgcatacatcaacacagtcgcttggggccctc 1320
|||||
1345 GCAGCTCAAGAACTAGAGAGAGCTGTCCAAAATGACAGAGGTTTACAGACATATGCTTCA 1404
|||||
1321 gcaagctcaagaactagaagagctgttcaaaatgacagaggttcaagacatagcttca 1380
|||||
1405 GATCAAGTTGACAAACAAATGGCTCATGATGATGATGATGATGATGATGATGATGAT 1464
|||||
1381 gatacagttcgaacaaatgagcctcatctgcttggggcccttcatcaagaataatgta 1440
|||||
1465 GGTGTCTCTCAGCGCTCCATCCAGCTTGAAGATTAAGATTAACCTCCAGAACACCCAG 1524
|||||
1441 gctgtctctcagcgctccatccagcttgagatgaaggttaacctccagaacacccag 1500
|||||
1525 TGGATGAATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1584
|||||
1501 tggatgaatgacagatgataatgataatgataatgataatgataatgataatgataatg 1560
|||||
1585 AACCTGGAACAAAGCAGCTTCCCAAAATCCCTTCTGGAATCCCAAGTGAACAAAGAGT 1644
|||||
1561 accctggacaaacagcagcctcccaaatccctcttgagatcccaatgagaaagcaaggt 1620
|||||
1645 GGCCTTGTAGTGAACAAACAAACAAATGACCTACTACCTCAAAATCCAGCATTTGATAG 1704
|||||
1621 ggccttgttagtgagacaaacacacaaatgacctactcccaatcccaagcatgtgtaag 1680
|||||
1705 GTTGGCACTTGGAAATACGTCTGACAGCAAGCTCAAAACCTTGACCTGACTGTACG 1764
|||||
1681 gtggcacttggaataatcagctgcaagcaagctcacaacacctgacactgactgcaag 1740
|||||
1765 TCCCGGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1824
|||||
1741 tcccggtgctcaaatgatacagctccctcccaatcagatgactcccaacagcaagagac 1800
|||||
1825 ACCAGCAAAATTTCCAGCCCTGTGATGATTTATGCAAAATTTCCGCAAGAGAGCTCCCA 1884
|||||
1801 accagcaaatccagccctctgtgattatgcaaatatctgccaagagagctcccca 1860
|||||
1885 ATTTCAGGCGCAGTGTCAACAGCCCTGATTTGATGATGATGATGATGATGATGATGAT 1944
|||||

Human C1CAL1 coding sequence, SEQ ID NO:4.
Human C1CAL1: goblet cell; mouse Gob-5 orthologue; drug screening;
expression inhibition; antisense therapy; gene therapy;
chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;

1465 GCTGTCCTCAGCGTCATCCAGCTTGAGAGTAAGGATTTAACCTCCAGAACAGCCAG 1524
 1441 gctgtcttcacgctccatccagcttgagagtaaggaattaacctccagaaacagccag 1500
 1525 TGGATTAATGGCAGTGTATCGTGACAGCAGCGTGGAAGAGACATTGTTCTTATC 1584
 1501 TTTTGTAAATGGCAGTGTATCGTGACAGCAGCGTGGAAGAGACATTGTTCTTATC 1560
 1585 ACCTGAGAACGACGCTCCCAATCTCTTGAGATCCCGTGGAGACAGAAAGGT 1644
 1561 acctggacaacgacgctcccaatccctctcttgagatcccgatggagaaagcaagt 1620
 1645 GGGTTGTAGTGACAAAAACACCAAAATGGCTACCTCCAAATCCAGGACATTGCTAAG 1704
 1621 ggccttgtagtgacaaaacacacaaatggcctaccatcccaagcattgtctaa 1680
 1705 GTTGGCAGCTTGGAAATATAGCTGTGACAGCAAGCTCAACAACTGACCTGTACAG 1764
 1681 gtggaacttggaataacagctgtcgaagcaagctcaaaccttgaccctgtactgacg 1740
 1765 TCCCGGCTGCAATGCTACCTGCTGCTCAATAGTGTGAGTGTGCAAAAGCAAGGAC 1824
 1741 tcccggtggtccaaatgctacccctgctcaatcaagtgactccaaagcaagac 1800
 1825 ACCAGCAATTCCTCCAGCCCTGTGTATGTTATGCAAAATTTGCGCAAGAGGCTCCCA 1884
 1801 accagaaatctcccaagccctgtgtatgtatgaataatctgcaagagagctccca 1860
 1885 ATTTCAGGCGCCAGTGTACAGCCCTGATTTGAAATGAGTGAAGGAAACAGTTACCTTG 1944
 1861 attctcagggcaggtgtcagccctgattgaatcagtgatgaataaacagttacttg 1920
 1945 GAACACTGATTAATGAGCAGAGTGTGATGCTACTAAGGATGAGGATGCTACTCAAG 2004
 1921 gaactactgtgataatgagcaggtgtctgtactcaagatgagcgtgtctctcaag 1980
 2005 TATTTCAACAATTATGACAGCAAGTGTAGATGATGATTAAGTGTGCGGCTGTGGAGGA 2064
 1981 tattcaacaattatgacagcaagtgatagatagatgaatgaggtgtcgtggagga 2040
 2065 GTTAAAGCAGCCAGAGGAGTGTATCCAGAGAGTGTGACAGCTGTACATCTGCG 2124
 2041 gttaacgacagcagcagcaggtgtatcccaagcagatgagcactgtacatactgagc 2100
 2125 TGGATTAATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2184
 2101 tggattgataatgataatgataatgataatgataatgataatgataatgataatgata 2160
 2185 GTTCAACACAGCAAGTGTGTTTACAGAGAACATCTCGGAGGCTCATTTTGTGCTTCT 2244
 2161 gtccaacacagcaagtgcttccagcagaacatctcgtggaggtccattgtgctctc 2220
 2245 GATGTCCCAAAATGCTCCATCTGATGCTCTTCCCACTGTGGCAATCACGACGCTGAAG 2304
 2221 gatgtcccaaaatgctccatctgactctctcccaactgtgcaaaatcacgacgctgaag 2280
 2305 GGGGAATTCACGGGGGAGTGTATTAATCTGACTTGAGCAGCTCTGGGAGTGAATTA 2364
 2281 ggggaattcacgggggagctcattatctgacttgagcagctcctgggaggtattat 2340
 2365 GACCATGGAACAGCTCACAAGTATATCATTTGAAATGAGTGAATGATGATGATGATGAT 2424
 2341 gaccatggaaacagctcacagatlatcatcgaataagtaagatattctgtgtcaga 2400
 2425 GACAAGTGAATGAATCTGTTCAAGTGAATGATGATGATGATGATGATGATGATGATGAT 2484
 2401 gacaagtgaatgaatctcttcaagtgataatcactgctctccaaaggaagcaaac 2460
 2485 TCTGAGGAGAGTCTTTTGTGTTAAACAGAAACATTTACTTTGAAATGAGCAGAGATCTT 2544
 2461 tctgaggagagctctttgtgttaaacagaaacatactctttgaaaaatggcagatctt 2520

2515 TTCAATGCTATTTAGGCTGTGTTAAGTTCGATTCGAAATATCCAAATATCCAACTTGA 2604
 2521 ttcatgtctatcagcgtgtgtatgaagtcgatctgaaatcagaataatccaaatgtga 2580
 2605 CGAGTATCTTTGTTTATTTCTCCACAGACTCCCGCAGAGACACCTAGTCTGATGAAAG 2664
 2581 cgagtatctttgtttatcttccacagactccgcagagacactgtctgtatgaagc 2640
 2665 TCTGCTGCTGTCTTATATTCATATTCATACAGACACCAATTCCTGGCATTCATTTAAA 2724
 2641 tctgctcctgtcttataatattatatacaagaacacatctcgtgcatcattttaa 2700
 2725 ATTATGTGAAGTGAATGAGAGAACTGACAGTGTCAATAGCC 2766
 2701 attatgtgaaagtgaatgagaaactgacagctgcatatagcc 2742

RESULT: 10
 AAH46120 standard; cDNA; 2843 BP.
 AAH46120:
 11-SEP-2001 (first entry)
 Mouse Gob-5 cDNA, SEQ ID NO:22.
 Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening; expression inhibition; antisense therapy; gene therapy; bronchial asthma; chronic obstructive pulmonary disease; antiasthmatic; ss.
 Mus sp.
 Key Location/Qualifiers
 CDS 15..2756
 FT /tag a
 FT /product- "Mouse Gob-5"
 W0200138530-A1.
 31 MAY 2001.
 22 NOV 2000; 2000MO-JP08232.
 24 NOV 1999; 99JP-0333479.
 27 APR 2000; 2000JP-0127589.
 (TAKE) TAKEDA CHEM IND LTD.
 Nakamishi A, Morita S;
 WPI: 2001-355935/37.
 P-ESDB: AAB73715.
 New antisense nucleotide, useful for treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease -
 Example 1; Page 89-91; 104pp; Japanese.

The invention relates to an antisense nucleotide targeted to the mouse Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart, the CLCA1 gene (coding sequence shown in AAH46102). The invention also relates to an antibody specific for the Gob-5 protein, medical and diagnostic compositions containing the antisense nucleotide or the antibody, and methods and kits for screening for compounds which inhibit the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells. The antisense oligonucleotides and antibody are therefore useful for the treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease. The present sequence represents mouse Gob-5 cDNA.

Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;

OY	385	CAGATGGGCAACCTGGGAGGAAAGGGGTGAAAGGATCCACCTCAGCCGATTTCAATCA	444
Db	368	catatgagacatgctgagagaaaaggggacacgagatccaactccctgactctttagca	427
OY	445	GGAAAAAGTTAGCTGGAATATGAGCACCACAGAGTAGGGCATTTGTCCATGAGTGGGCTCAT	504
Db	428	ggaagaagaactgactcagatctggccacaagaagaagaccttgcctacatgagtgggctcac	487
OY	505	CTACAGATGGGGAGATATTTGACGAGTGCACATTAATGATGACAAATTTCTACTTATCCAAATGA	564
Db	488	tcccgctggggagtggttataatgatacaacaagaagaaagttccactatccaagaagga	547
OY	565	AGAAATACAGCAGTAAAGATTTTCCAGCAGGTATTACTGTATCAAAATGTATGAAGAAGTGT	624
Db	548	aaacccaagaagcagtgagtgctcagacgacatccagcgtaaataatcaagttcgtgcgcg	607
OY	625	CAGGGAGGCAGCTGTTATCACCAA---AAGATGCACATTTAATTAAGTAACAGGACTCTAT	681
Db	608	caggagagccagtttctatcacttaacgggaagtgctgatatcgaagaagtaacggagctgatat	667
OY	682	GAAAAAGGATGTGATTTGTTTCTCCAAATCCCGCACAGCAGCAAGAGGCTTTCTAATGTGTT	741
Db	668	aaagacaatgctgtattcttgtaacagatccaacacaagaagagggcttcacatcgtctt	727
OY	742	GCACACATGTTGATTTCTATATGTTGAATTTCTGTACAGAACAAAACACACAAAGAGCT	801
Db	728	aaccaaatatcaattctcgtgctgtaattctgtacagaaaaataccaatacaagaagcgc	787
OY	802	CCAAACAGCAAAAATCAAAAATGCATATTCGCCAGACATGGAAGTGTATCCGTGATTTCT	861
Db	788	cccaatgaccaaaaaccaacgaatgcaactcccgaaagacagtgggaaagatcccaagaagatc	847
OY	862	GAGGACCTTTAAGAAACACACTCCTATGACAAACACACCCACCAAAATCCACCTTCATATG	921
Db	848	gaggactccaagcaaacacctcccaatgacagccagcgaactcgcacccaccttcctacgt	907
OY	922	CTGCAGATTGGACAAAGAAATTTGTGTGTTAAGTCCCTTGACAAATCTGGAAAGCATGGCGACT	981
Db	908	ctgcaaatgagacaagaatctgctgcttaagttcttgcgtataagtcggggagcagtcgtcgaac	967
OY	982	GGTAAACCGCCTCAATTCACATGGAATCAACAGGCGACACTTTTCTCGTCGCAGACACTTAA	1041
Db	968	gagatcgtcttctaaacgaaatgaatcgaagcgaagccgcttcttcgtcgaagcctgtagag	1027
OY	1042	CTGGGGCTTCCTGGGATGGGATGGTGAACATTTGACAGTGTCTCCCATATGTAAAGTAACTC	1101
Db	1028	caggagatccctgggtgcgggaatgctgacacttggacaagtgctcctatgataaagaagcaatc	1087
OY	1102	ATPACAGATTAACACTGTGCACATGACAGGAGACACTCTGCCAAAAAGATTACTCTCAGCAGCT	1161
Db	1088	aaaacagttaaacagtgctgctgtaacagagatctcgtataagaagacttaccacagtatct	1147
OY	1162	TCAGAGGAGGAGCTTCATCTGCACGGGGCTTCGATCGGCAATTTACTGTGATTAAGSAGAAA	1221
Db	1148	gcagagaggaacatctatatactcgtgccttcggacagcatttacagtgataaagaagaag	1207
OY	1222	TATCCACTGATGATGATGATGAAATTTGTGTGTCGTCGACGATGGGGAAGACAACTCTTAAT	1281
Db	1208	tatccaactgctggaatctggaatctgctgctgtaacgaaatgagggaagacaacacactagc	1267
OY	1282	GGGTGCTTTTAACGAGGTCAAAACAAAGTGGTGGCCATATCCACACAGTCCCTTTGGGGCC	1341
Db	1268	agctcgcttggcctcgtggaagcagagcgggggccatcatcatcagtggtgccttggacgc	1327
OY	1342	TCTGCACTCAAGAACTAAGAGAGCTGTCCAAAATGACAGAGGTTTACAGACATATGCT	1401
Db	1328	gctccgcctaaagaagccttgagcagctgctccaaaatgacaggaagccttgagacatactc	1387
OY	1402	TCAGATCAAGTTTTCAGAACATGGGCTCATATTGATGTGCTTTGGGGCCCTTTATCAGAGAAAT	1461
Db	1388	tcgagatcaggttcgaacaatgctctgttgaagcttctgcgacgactccctccaggaat	1447
OY	1462	GGAGCTGTCTCTCAGCGCTCCATCCAGCTTGAGAGTAAAGGATTAACCTCCAGAAACGC	1521

[illegible]

Db 2522 gatattcattgctatccagagctgtgataagtcacatcgaataatcccaac 2581
 QY 2599 ATTGCAGAGATCTTTGTTTATTCCTCCAGAGCTCCGACAGACACCTAGTCTGAT 2658
 Db 2582 attgcaagggtgtctgttctatcccccgcag-----gagcgcgcattcccgaa 2632
 QY 2659 GAAACGCTGCTCCTGCTTCTTAATTCATATACAGACACATTCCTGGCATTCAT 2718
 Db 2633 gactcaactccccctgtctcagcatcagcatcaacagaccattctgcatcacgtg 2692
 QY 2719 TTTAAATATATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2778
 Db 2693 ctgaagataatgctggaagctggaagctggaagctggaagctggaagctggaagct 2752
 QY 2779 TTTGTCAGATTAATAAATTAATTCATTCCTTTTGTGATTAATAAATTTTCAAA 2838
 Db 2753 ttcaagcaagaatacaacca-----gtcattccttcacacggaattttctaaaaa 2804
 QY 2839 TGTATTTTACCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2874
 Db 2805 tgaatttagactctcctgacgagcgatagtaa 2840

RESULT 12
 AAH46101
 ID AAH46101 standard; DNA; 2739 BP.
 AC AAH46101;
 XX 11-SEP-2001 (first entry)
 DE Mouse Gob-5 coding sequence, SEQ ID NO:3.
 XX
 KW Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
 KW expression inhibition; antisense therapy; gene therapy; bronchial asthma;
 KW chronic obstructive pulmonary disease; antisthmatic; ds.
 XX
 OS Mus sp.
 XX
 XX Key Location/Qualifiers
 FH 1.2739
 FT CDS
 FT /tag= a
 FT /partial
 FT /product= "Mouse Gob-5"
 FT /note= "No stop codon given in the specification"
 XX
 PN MO200138530-A1.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000MO-JP08232.
 XX
 PR 24-NOV-1999; 99JP-0333479.
 PR 27-APR-2000; 2000JP-0127589.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Nakanishi A, Morita S;
 DR WPI: 2001-355935/37.
 DR P-PSDB; AAB73715.
 XX
 PT New antisense nucleotide, useful for treatment and prevention of
 PT bronchial asthma and chronic obstructive pulmonary disease -
 XX
 PS Claim 3; Page 80-82; 104pp; Japanese.
 CC
 CC The invention relates to an antisense nucleotide targeted to the mouse
 CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
 CC relates to an antibody specific for the Gob-5 protein, medical and
 CC diagnostic compositions containing the antisense nucleotide or the
 CC antibody, and methods and kits for screening for compounds which inhibit

CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 CC This antisense oligonucleotide and antibody are therefore useful for the
 CC treatment and prevention of bronchial asthma and chronic obstructive
 CC pulmonary disease. The present sequence represents the mouse Gob-5
 CC gene coding sequence.
 XX
 XX Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other;
 XX
 XX Query Match 58 1%; Score 1733; DB 22; Length 2739;
 XX Best Local Similarity 78.0%; Pred. No. 0;
 XX Matches 2145; Conservative 0; Mismatches 585; Indels 21; Gaps 4;
 XX
 QY 15 ATGGGCGCCATTTAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 84
 Db 1 atggaatcttgaagagctgtctcctccttgcattccctcctcctcctcctcctcctcctc 60
 QY 15 AGTAATTCACCTATTCAGCTGGAACAAATGCTATGAAAGGCAATGCTGCTGCTGCTGCTG 144
 Db 1 atggaatcttgaagagctgtctcctcctcctcctcctcctcctcctcctcctcctcctc 120
 QY 15 CCCAATGTCGCGAAGATGAAACATCATTCMAAATAAAGAGACATGAGACCCAGCGCA 204
 Db 1 cagcagctgacggaagatgaagccctcattcaacataaagacatggtgactcagcgc 180
 QY 205 TCTGTATCTGTTTGAACCTACAGAGAGCGATTTTATTTAAATGTTGCCATTTTG 264
 Db 1 tctccatcctgttgaagctgaagctgaagctgaagctgaagctgaagctgaagctgaagct 240
 QY 285 ATTCTGAAACATGAG 324
 Db 1 attcccgagctggaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 300
 QY 385 AATCGATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 384
 Db 1 aacgctgactcctgtatcaacaacacccctcgaagatgaagcctcgaagcctcgaagcct 360
 QY 385 CAGATGGCCAACTGTGAG 444
 Db 1 catatagagcagctgtgag 420
 QY 445 GGAATAAAGCTTACCTGAATATGAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 504
 Db 1 ggaataaagctgtgactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 480
 QY 505 CTACGATGGGAGTATTTGACGAGTACATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 564
 Db 1 ttccgaltgagagctgttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 540
 QY 505 AGAATACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 624
 Db 1 aaaccccaagcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 600
 QY 625 CAGGAGAGAGCTGTTACACCAA---ACATGCACTTCAATTAAGTAACAGAGCTGAT 681
 Db 1 caggagagcagtgatcctaacaggaagtgatgaacagcagcagcagcagcagcagcagcag 660
 QY 682 GAAAAAGATGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 741
 Db 1 aaagcaaatgtgtatctgtatccagatccacacaaacaggaagcgttcacatcagtggtt 720
 QY 722 GCACAACATGTTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
 Db 1 aaccaaaatataatctgt 780
 QY 802 CCAACAGAGCAAAATCAAAATGCAATCTCCAGAGCAATGGAAGTATGATCTGCTGATCT 861
 Db 1 ccaaatgacaaacaaacagatgtcaatctccagagacgtgtggaagctacacagagc 840
 QY 862 GAGGATTTAAGAAACCTCTGATGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921
 Db 1 gaggaactcaagaac 900

QY	922	CTCGAGATTGGACAAACAAATTGGTGTGTTTAAAGTCCTTGCACAAATTCGGAAGTCGGCCACT	981
Db	901	ctgcaaatlyggaacaagaattgctgcttaagtlcttgataaagtcgggaagcatgctgtaac	960
QY	982	GGTAACGCGCTCAATGCATGCAATCAAGACGGCCAGCTTTTCTGCTGCAGACACTTGAG	1041
Db	961	gatgatcgtcttaaccgaatgaatacgaagcaagccgtcttctcgtcgtcgaagctgtag	1020
QY	1042	CTGGGGTCTCTGGGTTGGGATGGTGACATTTTGACACAGTGCCTCCCATGTACAAAGTAATC	1101
Db	1021	caaggatctccggtgacggaatcgtgaccttctgaagcgtcgtcctatgatacaagcgaatc	1080
QY	1102	ATPACGATMAAACAGTGGCAGCTGACAGGACACACTGCGCAAAAGATTTACTGACACACT	1161
Db	1081	aaacaaglttaaacagltggtgctgcaagagatcgtcgtataaagcacttaaccacaglatct	1140
QY	1162	TCAGAGAGGAGCTGCATCTGCACGGGCTTCGATCGCATTTACTGTGATTAGAAAGAA	1221
Db	1141	gcagagaggagacatctatatgctcgtccttcggaacagcatltaacagtgataaagagag	1200
QY	1222	TATCCAACTGATGGATGTGAATTTGTGTGCTGACGGATGGGGAAGACACACTTAAAGT	1281
Db	1201	tatccaactgatgatactgaaatctgacgtcgtcgtacagcgaatgagggaagacaacacatlagc	1260
QY	1282	GGGTGCTTTTAAACGGGCAAAACAAAGCGGCCATATCCACACAGTGCCTTTGGGGGCC	1341
Db	1261	agctgccttcaacctgctgtaagcagagccaggacatccaatcaacagtgagccctggaagcg	1320
QY	1342	TCTGCAGCTCAGAGACTAGAGAGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCT	1401
Db	1321	gctgcgctaaagagcctgtagcagcgtgtccaaaatgacagagagcgtcgtcgaacatacct	1380
QY	1402	TCGATCAAGTTGACGAACATGGCCTCATTTGATGCTTTTGGGGCCCTTTTCATCAGGAAT	1461
Db	1381	tcgagatcaaglttcagaaacaaatgctctgtatgacgtcttcgcagcaactctccacgaagaat	1440
QY	1462	GGAGCTGTCTCTCGAGCCTCCATCCAGGCTTGAAGTAAGTAAGATTTAACCCCTCCAGAAAGC	1521
Db	1441	gcggcgatcgctcagcaactccacccagcctggaagcagagggatlaactccagaaataac	1500
QY	1522	CAGTGGATGATGGCACAGTGAATCGTGGACAGCAGCGTGGGAAAGACACTTGTGTTCT	1581
Db	1501	caatgtagtgaatggtctcagtgatcgtggaagctggtggaagagacaactgttctt	1560
QY	1582	ATCACTGTGACACGACGCTCCCAATCTTCTTGCGGATCCCAATGTGGACAGACGA	1641
Db	1561	atcacctggaacaagcactccctccaatataltatctggaatcccaagcgaagtgaacaa	1620
QY	1642	GGTGGCTTTGATGGCAAAACACCAAAAGGCCACTCTCCAAATCCGACGGCACTTGCT	1701
Db	1621	aatggtlltlaaacatgaagcaaacacaaatgaagtggtcctaccctcaagtlcccaagcaggtc	1680
QY	1702	AAGTTGGCACTTGGAAATACAGTGTGCAGAGAGCTCCAAACCTTGACCTGTGACTTC	1761
Db	1681	aaggttgctctttggaatatcagcatcaacgaagagctcagactcctcaacttgactgtc	1740
QY	1762	ACGTCCGTGCTCAATGCTACCTGCTCTCCAAATTACAGTGACTTCCAAAACGACAAAG	1821
Db	1741	acctcccgtaggaagtgctacacagctcctcatltaacagtagaccocgtagtagataag	1800
QY	1822	GACACGACGAAATTTCCCAAGCCTCTGTGATGTTATGCAAAATTTTCCGACAGACCTCC	1881
Db	1801	aacacaggaatattcccaagcctctgtaaacagtgtatgcaagatctgcgaagaagacctcg	1860
QY	1882	CCAATTTTCAGGGCAGTGTGCACAGCCCTGATGTGATGATGCAATGGAAGAAAACACTTTC	1941
Db	1861	ccattctcgaagcgctgcacaagccttgatgtaactcgtgaaatggaataacagtaaac	1920
QY	1942	TTGGAACACTAGTGAATATGGAGCAGGTGTGTGATGATCTACTAAGATGACGGTGTACTACA	2001
Db	1921	ctggaatlaactgagtaacagcaggtgtcgaatgccaacaagaatgagtgtctactca	1980
QY	2002	AGGATTTTACACACTTATGACACGAATGATGATACAGTGTAAAGTCCGGCTCTGGCA	2061

Db	1991	aggttttttaacagcttcttgatgcaaatggtgagtaacagtgtaaatatagggctctggga	2040
Gy	2052	GGAGTTAACGACCGACAGCGAGAGTGTATACCCGACAGAGTGGACACTGTATACCT	2121
Db	2011	ggagtcacttaagacagacagagagagcagcactccgaagaacagagcccatgtaactagat	2100
Gy	2112	GGCGGANTGTGAGTATGATGAAATATACAAATGGAAATCCACCAAGACCTCCAAATTAATAGAGT	2181
Db	2101	ggcttggatcttgagagcttgtagaataagaatgaaaccacacagctcttgaacta-----gt	2154
Gy	2112	GATGTTTAAACACAAAGCAAGTGTGTTTTCAGCAGAAATCCCTGGGAGGCTCATTTGTGGCT	2241
Db	2155	tatgttcaagacaaagacagcgtgtgtcttcagagaacactcttcaggggagtcgtttgagc	2214
Gy	2242	TCTGATGTC--CAAATGCTCCCATACCTGATCTCTTCCACCTGGCCAAATACCGAC	2296
Db	2215	acaatgtgccccgacagagctctccatctcgtacctctcttccacacctgtcaactatgac	2277
Gy	2239	CTGAGAGCGGGAATTTCCAGGGGGGACGTCCTATTAATCTGACTTTGGACACACTCTCTGGGAT	2355
Db	2215	ctgaagagccagatcccaagagcagaacctggtagaactgaactgagcttgccttcgggagat	2334
Gy	2339	GATTATGACCATGGAGACAGCTCCAACTATATCATTCATTCGAAATAGTCAAGTATATCTTGAT	2418
Db	2335	gactacgacccacgggagagagcttccaactatacatccgaaatgagcaccagtatctgtag	2399
Gy	2419	CTCAGAGACAAGTTTCAATGATCTCTTCAAGTGAATACTACTGCTCTCTCAATCCCAAGGA	2478
Db	2335	ctcagggagccacttcaacacctcaactccaagtgaacactacccgctcttaaccacaagag	2454
Gy	2419	GCCACACTCTGAGAGAGCTTTTGTGTTTAAACCGAAGAAACATTACTTTTAAATATGGACA	2538
Db	2435	gcccagctctctgggaaactcttcttgattgtaacctggaggagaaaccttcttggaaatggcaca	2514
Gy	2539	GATCTTTTCATTGCTTATTCACGGCTGTGATTAAGAGTGGATCGATCGAATCAGAATATTCAC	2598
Db	2515	gatatcttcaatctgtaaccagcgctgtggataaagatccaatctgaaatcagaataatctccaac	2578
Gy	2539	ATTGGACGAGTATCTTTGTTTATTCCTCCACAGACTCCGCGACAGACACTATGCTTGAT	2658
Db	2535	atggaagcgggtctgtgttcaatccccgcctcag-----gagcgcccatctccgaa	2625
Gy	2619	GAACGCTGCTGCTCTGTCTCTAATATTCATATTAACAAGACCACTTCTGGCATTTACATT	2718
Db	2636	gactcaatctcccccctgtctctgcatacagatcaacagacacatctctgtgcatccaagtg	2685
Gy	2719	TTTAAAAATTATGTGAGAGTGGATAGGAGACTCCAGCTGCAATAGCCAG	2769
Db	2636	ctgaagaataatgttgaagatggtctagaaggaaatgcagatgaaactagatttg	2736
RESULT 13			
HAZ65G93			
Db	AA:65095	standard; cDNA; 3265 BP.	
Gy	AA:65095;		
Db	05-APR-2000	(first entry)	
Gy	Membrane-bound protein PRO1124	encoding cDNA.	
Db	Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;		
Gy	pharmaceutical; receptor immunoadhesin; gene mapping; ss.		
Db	HeMo sapiens.		
Gy	HeMo sapiens.		
Db	WO9963088-A2.		
Gy	WO9963088-A2.		
Db	09-DEC-1999.		
Gy	09-DEC-1999.		
Db	02-JUN-1999;	99MO-US12252.	
Gy	02-JUN-1999;	99MO-US12252.	

PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091489.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 10-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.
PR 19-AUG-1998; 98US-0097218.
PR 20-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098525.
PR 16-SEP-1998; 98US-0100634.
PR 12-JAN-1999; 99US-0115565.
PR (GENE) GENENTECH INC.
PR Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PR Wodt WI, Yuan J;
PR WPI: 2000-072883/06.
PR P-SDS: AA166749.
PR Membrane-bound proteins and related nucleotide sequences -

Db 1818 taaatgataagagcgaacagcttcccccagcccaagatgcttaccagcagaatctt 1877
 1869 CCAGGACGCTCCCAATTCAGGCGCAGTGTACAGCCGCTGATGTAATGTAATG 1928
 1878 acaagagatactactgcttcttggaagcccaatgactgcttccatgatacagaatg 1937
 1929 AAAACAGTACCTTGGAACTAGTGAATGAGGAGGCTGTGATGCTACTAAGATGA 1988
 1938 acatacagaagcttggacacttcttgataatgctgacgctgcttcttcaagaatga 1997
 1969 CGGTCTACTACTAAGTATTTACAACTTATGACAGGAATGATGATGATGATGAT 2048
 1998 tggagcttactcagaagcttcttaccagcatalacagaataatggcagataagctlaaagt 2057
 2049 GGGGCTCTGGAGAGGTTAAGCGACGAGAGAGTGAATGATGATGATGATGATGATG 2108
 2058 tggagcttactcagaagcacaacacgctcgaagcttcaagctcgaagctcgaagctcga 2117
 2109 ACTGATACCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2168
 2118 cgcgtacatacagcagctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2177
 2169 AATTAATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2228
 2178 aatgat---gaggaactcagacacacacacacacacacacacacacacacacacacac 2234
 2229 CTGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2288
 2235 tgcatttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2294
 2289 AATCAGCAGCTGAAAGCGGAAATTCAGCGGGGAGTCTGATTAATGATGATGATGATG 2348
 2295 aatcacagac 2351
 2349 TCCCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2408
 2352 accagagataatcttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2411
 2409 TATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2468
 2412 tatcttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2471
 2469 CCAAAAGAGCAACTGAGGAAAGTCTTTGTTAAACAGAAACATTAATGATGATGATG 2528
 2472 accaagagagcgaactccaagaaagcttgcatttaaccagaaataatctcagaaga 2531
 2529 AATGGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2588
 2532 aaatgac 2591
 2589 AATATCAACATGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2645
 2592 agtaccac 2651
 2646 ACCTAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2683
 2652 tcttaccactactcctactcctactcctactcctactcctactcctactcctactcct 2689

Query: Match 43.7%; Score 1304; DB 22; Length 3265;
 Best Local Similarity 69.8%; Pred. No. 0;
 Matches 1869; Conservative 0; Mismatches 785; Indels 24; Gaps 7;
 31 CCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 140
 31 ---aaatacttctcatttaagctgataataatgcttgaagataatgctatgctat 137
 141 CGACCCCAATGTCGAGAGAGATGAACACTCATTTCAACAAATGAAGACATGTCGACCA 200
 141 ---aaatacttctcatttaagctgataataatgcttgaagataatgctatgctat 137
 198 agcttctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 257
 201 GGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 260
 261 TTTGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 320
 261 ---aaatacttctcatttaagctgataataatgcttgaagataatgctatgctat 137
 268 attaatctcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 317
 321 CAAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 380

OY	2529	AAATGGACAGACTTTTCTTTCATTCAGGCTGTGATTAAGTCGATTCGAATGAGA	2589		
Db	2532	aaatgcacaacccatatttctttgcatataaagtagataaagaacattgcacaa	2591		
OY	2589	AATATCCACACATTGCACGAGTATCTTTGTTTATTCCTCCACAGACTCCGACAGA--C	2645		
Db	2592	agataccaacattgcaacaagtaacttgitlatacctcaagaacactctgtgacattga	2651		
OY	2646	ACCTAGTCCCTGAGGAACACGCTCCTCTGTGCTTAATA	2683		
Db	2652	tctctacacctactctactctactctactctactgata	2689		
RESULT 15					
AAFA44241					
ID	AAFA44241	standard; cDNA; 3265 BP.			
XX	AAFA44241;				
AC					
XX					
DT	02-Apr-2001	(first entry)			
XX					
DE	Human PRO1124 (UNO562)	nucleotide sequence SEQ ID NO:378.			
XX					
KW	Human; secreted and transmembrane protein; PRO; cytosolic;				
KW	cell death; cancer; chromosomal mapping; gene mapping; tissue typing;				
FW	diagnostic assay; ss.				
XX					
OS	Homo sapiens.				
XX					
PN	WO200073454-A1.				
XX					
PD	07-DEC-2000.				
XX					
PF	30-MAR-2000; 2000MO-US08439.				
XX					
PR	02-JUN-1999; 99MO-US12252.				
PR	23-JUN-1999; 99US-0141037.				
PR	07-JUL-1999; 99US-0143048.				
PR	20-JUL-1999; 99US-0144758.				
PR	26-JUL-1999; 99US-0145698.				
PR	28-JUL-1999; 99US-0146322.				
PR	17-AUG-1999; 99US-0149396.				
PR	15-SEP-1999; 99MO-US21090.				
PR	15-SEP-1999; 99MO-US21547.				
PR	08-OCT-1999; 99US-0158663.				
PR	30-NOV-1999; 99MO-US28313.				
PR	01-DEC-1999; 99MO-US28301.				
PR	16-DEC-1999; 99MO-US30095.				
PR	20-DEC-1999; 99MO-US30911.				
PR	05-JAN-2000; 2000MO-US00219.				
PR	06-JAN-2000; 2000MO-US00376.				
PR	11-FEB-2000; 2000MO-US03565.				
PR	18-FEB-2000; 2000MO-US04341.				
PR	22-FEB-2000; 2000MO-US04414.				
PR	24-FEB-2000; 2000MO-US04914.				
PR	24-FEB-2000; 2000MO-US05004.				
PR	02-MAR-2000; 2000MO-US05841.				
PR	15-MAR-2000; 2000MO-US06884.				
PR	20-MAR-2000; 2000MO-US07577.				
XX					
PA	(GETH) GENENTECH INC.				
XX					
PI	Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;				
PI	Ferrera NJ, Fong S, Geider H, Gerritsen ME, Goddard A, Godowski PJ				
PI	Grimaldi CJ, Gurey AL, Kijavini IJ, Napier MA, Pan J, Paoni NF;				
PI	Koy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;				
XX	Zhang Z;				
XX					
XX	WPI; 2001-032160/04.				
DR	P-PSDB; AAB65272.				
XX					
PT	PRO polynucleotides used to produce polypeptides used to target				

biactive molecules such as toxins, radiolabels or antibodies, to
specific cells, to cause targeted cell death -

Claim 2, Fig 273, 935pp: English.

The present invention describes human secreted and transmembrane PRO
proteins. The PRO proteins have cytostatic activity. The PRO proteins
can be used for targeted delivery of bioactive molecules, such as
toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
sequences, and their fragments, can be used as hybridization probes, in
chromosomal and gene mapping, and in the generation of anti-sense RNA
and DNA. They may also be used to produce transgenic animals which are
used to develop and screen therapeutically useful reagents. The PRO
nucleotide and protein sequence can be used for tissue typing and in
treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
AA144270 to AA144470 represent PCR primers and hybridization probes used
in the isolation of human PRO sequences. AA144087 to AA144265 and
AA1465154 to AA1465300 represent human PRO polynucleotide and protein
sequences given in the exemplification of the present invention.

Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;

Query Match	43.7%	Score 1304;	DB 22;	Length 3265;
Best Local Similarity	69.8%	Pred. No. 0;		
Matches 1869;	Conservative	0;	Mismatches 785;	Indels 24; Gaps 7

[illegible]

D 678 gtagaagaagatgttcaattcttccgtataaagtaacaaacgaagaaagatccataat 737
Q 738 GTTGGCAACACATGTGATCTATAGTTGAATCTGTACGAACAAACCAACACAAAGA 797
D 738 gttatgtaagaagatgtatctgtctgtatgtatgtatgaacaaacacccaatacaaga 797
Q 798 AGCTCAACAAAGCAACAAATCAAAATCAATCTCCGAAGCAAGCAAGGAGATCCGTGA 857
D 798 agctcaacgaagctacaaacataaagtgcaattttagaagatcaatggagagtgatga 857
Q 858 TTCTGAGAGCTTTAAGAAACCACTCCTATGACACACAGCCACCAATCCACTTCTC 917
D 858 ttctgagatattttaaaccacacacacacacacacacacacacacacacacacacac 917
Q 918 ATTGCTGACATGTGACAAAGATTTGTGTAGTCTGCTTGTGACAAATCTGGAACATG 977
D 918 attgtgaaagctcagcaaaacatgtgtctgtatgtatgtatgtatgtatgtatgtatgt 977
Q 978 GACTGTACACCGCTCAATGACTGATCAACAGCAGCCAGCTTTCTGTGACAGACGT 1037
D 978 ggttaagagacgcctaaatcgaatgaatgaacgaacaaacacacacacacacacacac 1037
Q 1038 TGACCTGGGCTCTGGGTGGGATGGTGACATTTGACAGTCTCCATGACAAAGTGA 1097
D 1038 tgaacatggaatccctgt 1097
Q 1098 ACTGATACAGATTAACAGTGCAGTGCAGACAGGACACACTGCCAAAGATTACCTGAC 1157
D 1098 gctaatcaataaagaacagtgatgaagaagaacacacacacacacacacacacacacac 1157
Q 1158 AGCTTACGAGGAGGAGTCCATCTGACGCGGCTTCGATCGGATTTACTGTATTAGAA 1217
D 1158 tccctgtggaagaaatccatctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1217
Q 1218 GAATAT---CCAACTGATGATCTGAATTTGTGCTGTGACGATGGGGAAGACACAC 1274
D 1218 gctaatcccaactcgaatggaatggaatggaatggaatggaatggaatggaatggaatgga 1274
Q 1275 TATAGTGGGCTTTAAACAGAGTCAACAAAGTGTGCTCAATCCACAGACAGTCTT 1334
D 1275 tgaagatctgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgt 1334
Q 1335 GGGGCGCTCTGACGCTCAACAGTGAAGAGAGCTGTCCAAATGACAGAGGTTTACAGAC 1394
D 1335 ggggaagagctgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgt 1394
Q 1395 ATATGCTTCAATCAAGTTCGAACAAATGCTCATGATGATCTTTGGGCGCTTTCATC 1454
D 1395 ttaatgttcaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgta 1454
Q 1455 AGGAATGGAGCTGTCTCTACGCGTCCATCCAGCTTGAAGTGAAGGATTACCTTCA 1514
D 1455 aggaatgagctgtctctacgctgtctctacgctgtctctacgctgtctctacgctgtctct 1514
Q 1515 GAACAGCCAGTGTGAATGTCACAGTGTGTCGACAGCAGCTGGGGAAGGACACTTT 1574
D 1515 tagtaatgctgtgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgt 1574
Q 1575 GTTCTTATCTGACGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1634
D 1575 ctctctatcaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgta 1634
Q 1635 GAACCAAGGCTTTGTAGTGAACAAACCAACCAACCAACCAACCAACCAACCAACCA 1694
D 1635 aataatggaatattcacagtgatgtacacacacacacacacacacacacacacacacacac 1694
Q 1695 CATGCTAAGTGTGACATTTGGAATACAGTCTG-----AAGCAAGCTCACAAACCTT 1748
D 1695 aactgcaaaagtgagac 1748
Q 1749 GACCTGACGTCAGTCCGCTGCGTCAATGCTACCTGCTCCAAATTAAGTGAATTC 1808
D 1749 aactatcaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgta 1808
Q 1758 aactatcaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgta 1817

Q 1809 CAAACGACAGGAGACACGACAAATTCGCCAGCCCTGTGTAGTTATGCAATATTCG 1868
D 1809 taaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgta 1868
Q 1869 CCAAGAGCTCCCAATTTCTGACGCGCAGTGTACAGCCCTGATGATGATGATGATGATG 1928
D 1869 acaagagtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgta 1928
Q 1929 AAAACAGTACCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1988
D 1929 acaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgta 1988
Q 1989 CGGTGTACTCAAGTATTTCAACTATTTGACAGCAAGTATGATGATGATGATGATGATGATG 2048
D 1989 tgaagatctatcaagatatttataagataataagataataagataataagataataagataata 2048
Q 2049 GCGGCTCTGAGAGAGTAAACGACGACGACGACGACGACGACGACGACGACGACGACG 2108
D 2049 tgggctcagtgagagac 2108
Q 2109 ACTGATACCTGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2168
D 2109 cggctacatccagctgt 2168
Q 2169 AATTATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2228
D 2169 aattgat---gagatatacagac 2228
Q 2229 CTGATTTGTGCTTCTGATGCTCCAAATGCTCCATGCTGATGCTGCTCCACTGCGCA 2288
D 2229 tgcattgt 2288
Q 2295 AATCAGCAGCTGAAAGCGGAAATTCACGGGAGCTGCTGATGATGATGATGATGATGATG 2348
D 2295 aatcagcagctgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtat 2348
Q 2349 TCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2408
D 2349 accagagataatattgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtat 2408
Q 2409 TATCTTGTATCTGACAGACAGTCAATGATGATGATGATGATGATGATGATGATGATG 2468
D 2409 tattctgtatctaaagagacagctgtatgtatgtatgtatgtatgtatgtatgtatgtatgtat 2468
Q 2472 CCCAAGGAGGACCACTGAGAGCTTTTGTGTTAAACCAACCAACCAATTAATCTTTGA 2528
D 2472 accaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2528
Q 2529 AATATCAACATGTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2588
D 2529 aatatacaatgtgcaaaagaaactgtgtatcccccacaaatccctgtatgtatgtatgtatgt 2588
Q 2532 AACTAGTCTGATGAACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2683
D 2532 tccatacactactcactcactcactcactcactcactcactcactcactcactcactcactcact 2683
Q 2683 AACTAGTCTGATGAACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2683
D 2683 tccatacactactcactcactcactcactcactcactcactcactcactcactcactcactcact 2683

Search completed: April 4, 2002, 03:25:02
Job time: 60654 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:16:12 ; Search time 22700.8 Seconds

(without alignments)
2167.812 Million cell updates/sec

Title: US-09-049-696-20

Perfect score: 2983
Sequence: 1 GAAATCAGCAGGAGATGAC.....AAATGCTAAACCAACTGGCTA 2993

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenEmbl.*
1:  gb_da:*
2:  gb_da:*
3:  gb_da:*
4:  gb_da:*
5:  gb_da:*
6:  gb_da:*
7:  gb_da:*
8:  gb_da:*
9:  gb_da:*
10: gb_da:*
11: gb_da:*
12: gb_da:*
13: gb_da:*
14: gb_da:*
15: gb_da:*
16: gb_da:*
17: gb_da:*
18: gb_da:*
19: gb_da:*
20: gb_da:*
21: gb_da:*
22: gb_da:*
23: gb_da:*
24: gb_da:*
25: gb_da:*
26: gb_da:*
27: gb_da:*
28: gb_da:*
29: gb_da:*
30: gb_da:*
31: gb_da:*
32: gb_da:*
33: gb_da:*
34: gb_da:*
35: gb_da:*
36: gb_da:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2966.2	99.4	3311	6	AX193489	AX193489 Sequence
2	2966.2	99.4	3311	6	AF039400	AF039400 Homo sapi
3	2806.8	94.1	2826	9	AF127036	AF127036 Homo sapi
4	1999.4	67.0	2022	9	AK024970	AK024970 Homo sapi
5	1924.6	64.5	3079	4	AF095584	AF095584 Sus scrofa
6	1764	59.1	2937	10	AB017156	AB017156 Mus muscu
7	1310.2	43.9	3204	9	AF127035	AF127035 Homo sapi
8	1304	43.7	3221	9	AK000072	AK000072 Homo sapi
9	1304	43.7	3265	6	AX092338	AX092338 Sequence
10	933.4	31.3	2984	4	AF036445	AF036445 Bos tauru
11	900.2	30.2	3317	4	AF001261	AF001261 Bos tauru
12	865.4	29.0	2765	10	AF115852	AF115852 Mus muscu
13	865.4	29.0	3137	10	BC008147	BC008147 Mus muscu
14	863.8	29.0	3058	10	AF108501	AF108501 Mus muscu
15	863	28.9	3022	10	AF047838	AF047838 Mus muscu
16	863	28.9	3471	10	AF052746	AF052746 Mus muscu
17	831	27.9	3415	9	AF043976	AF043976 Homo sapi
18	806	27.0	2820	4	AF001263	AF001263 Bos tauru
19	800.6	26.8	3288	4	AF001262	AF001262 Bos tauru
20	790.8	26.5	878	6	I95746	I95746 Sequence 8
21	672.2	22.5	1895	9	AK000138	AK000138 Homo sapi
22	608	20.4	113764	9	HSJ651E10	AL122002 Human DNA
23	608	20.4	164891	2	AL358950	AL358950 Homo sapi
24	594.4	19.9	35278	9	AF039401	AF039401 Homo sapi
25	553.8	18.6	2832	6	AF127980	AF127980 Homo sapi
26	553.8	18.6	4077	6	AX054697	AX054697 Sequence
27	552.2	18.5	2970	9	AF043977	AF043977 Homo sapi
28	550.6	18.5	3604	9	AB026833	AB026833 Homo sapi
29	381	12.8	1265	4	AF001264	AF001264 Bos tauru
30	323.8	10.9	401	6	AX192467	AX192467 Sequence
31	187	6.3	124096	2	AL354988	AL354988 Homo sapi
32	163	5.5	140718	2	AL356270	AL356270 Homo sapi
33	151.6	5.1	375	6	AX071171	AX071171 Sequence
34	131.8	4.4	140718	2	AL356270	AL356270 Homo sapi
35	108.4	3.6	175591	2	AC068071	AC068071 Homo sapi
36	91.2	3.1	350	6	AX188011	AX188011 Sequence
37	91.2	3.1	454	6	AX185116	AX185116 Sequence
38	88	3.0	549	6	AX054715	AX054715 Sequence
39	73	2.4	307	6	AX187657	AX187657 Sequence
40	73	2.4	333	6	AX186584	AX186584 Sequence
41	73	2.4	423	6	AX184779	AX184779 Sequence
42	68	2.3	227	6	AX186451	AX186451 Sequence
43	52	1.7	7218	6	I66494	I66494 Sequence 14
44	49.4	1.7	1141	6	AX083744	AX083744 Sequence
45	47.8	1.6	110	10	AF077303	AF077303 Rattus no

ALIGNMENTS

RESULT 1
AX193489 AX193489 3311 bp DNA PAT 15-AUG-2001

DEFINITION Sequence 1056 from Patent WO0149716.

ACCESSION AX193489

VERSION AX193489.1 GI:15211440

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

FEATURES

schrice

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 3311)
Xu, D., Lodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J.,
Stolk, J.A., King, G.E., Wang, T. and Jiang, Y.

Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use

Patent: WO 0149716-A 1056 12-JUL-2001;
CORIXA CORPORATION (US)

Location/Qualifiers
1..3311

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 99.4%; Score 2966.2; DB 6; Length 3311;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2979; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 GAATTCACAGGAGATGTACAGCAATGGGCCATTTAAGAATTCTGTCTCATCTTGAATT 60
DB 328 GGAATTCACAGGAGATGTACAGCAATGGGCCATTTAAGAATTCTGTCTCATCTTGAATT 387
QY 61 CTTCACCTCTAGAGAGGGGCCCTGAGTAATTCACATTCAGCTGGAACAACAATGGCAT 120
DB 388 CTTCACCTCTAGAGAGGGGCCCTGAGTAATTCACATTCAGCTGGAACAACAATGGCAT 447
QY 121 GAAGCATTTGCTGTGCAATGACCCCAATGTGCCAGAAATGAAACATCATTCACAA 180
DB 448 GAAGCATTTGCTGTGCAATGACCCCAATGTGCCAGAAATGAAACATCATTCACAA 507
QY 181 ATAAAGCATGTGTGACCCAGCATCTCTATCTCTTTGAAGCTACAGAAAGCATTT 240
DB 508 ATAAAGCATGTGTGACCCAGCATCTCTATCTCTTTGAAGCTACAGAAAGCATTT 567
QY 241 TATTTCAAAATGTTGCCATTTTGAATTCCTGAACATGGAAGACAAAGCTGATGTG 300
DB 568 TATTTCAAAATGTTGCCATTTTGAATTCCTGAACATGGAAGACAAAGCTGATGTG 627
QY 301 AGACCAAACTGTGACCTACAAAAATGCTGATGTTCTGTTGCTGATCTACTCTCA 360
DB 628 AGACCAAACTGTGACCTACAAAAATGCTGATGTTCTGTTGCTGATCTACTCTCA 687
QY 361 GGTATGTATGAACCTTACATGAGATGAGCACTGTGAGAAAGGAAAGGAAATG 420
DB 688 GGTATGTATGAACCTTACATGAGATGAGCACTGTGAGAAAGGAAAGGAAATG 747
QY 421 CACCTCACTCTGATTTGACAGAAAGGTTAGCTGAATATGACCAAGAGTATG 480
DB 748 CACCTCACTCTGATTTGACAGAAAGGTTAGCTGAATATGACCAAGAGTATG 807
QY 481 GCATTTGTCATGAGTGGCTCATCTACGATGGGAGTATTTGACGAGTACATATGAT 540
DB 808 GCATTTGTCATGAGTGGCTCATCTACGATGGGAGTATTTGACGAGTACATATGAT 867
QY 541 GAGAAATCTCTATTCGAATGGAAGATPACAGCATAGATGTTGACAGAGTATCT 600
DB 868 GAGAAATCTCTATTCGAATGGAAGATPACAGCATAGATGTTGACAGAGTATCT 927
QY 601 GGTACAAATGTATGAAGAGTGTGAGGAGGAGCAGCTGTACCAAAAAGATGACATTC 660
DB 928 GGTACAAATGTATGAAGAGTGTGAGGAGGAGCAGCTGTGTACCAAAAAGATGACATTC 987
QY 661 AATAAGTACAGAGCTCTATGAAAAAGATGTAGTTTGTCTCCATCCCGCCAGAG 720
DB 988 AATAAGTACAGAGCTCTATGAAAAAGATGTAGTTTGTCTCCATCCCGCCAGAG 1047
QY 721 GAGAGGCTCTATATGTTGACACATGTTGATTCATATGATGAAATTCGTACAGA 780
DB 1048 GAGAGGCTCTATATGTTGACACATGTTGATTCATATGATGAAATTCGTACAGA 1107
QY 781 CAAAACCAACAAAGAGCTCCAAACAGCAAAATCAAAAATGCAATCTCCGAAGCA 840
DB 1108 CAAAACCAACAAAGAGCTCCAAACAGCAAAATCAAAAATGCAATCTCCGAAGCA 1167
QY 841 TGGGAAGTATCCGTGATTTCTGAGACATTTAAGAAACCACTCTATGACAAACAGCA 900
DB 1168 TGGGAAGTATCCGTGATTTCTGAGACATTTAAGAAACCACTCTATGACAAACAGCA 1227
QY 901 CCAATTCACACCTTCTCATTTGCTGAGATGAGCAAAAGATGTGTGTTAGTCTTGAC 960
DB 1228 CCAATTCACACCTTCTCATTTGCTGAGATGAGCAAAAGATGTGTGTTAGTCTTGAC 1287

QY 961 AATCTGGAAACATGGGCACTGTATACCGCTCAATGACTGATCAAGCAGGCAGCTT 1020
DB 1288 AATCTGGAAACATGGGCACTGTATACCGCTCAATGACTGATCAAGCAGGCAGCTT 1347
QY 1021 TTCTGCTGCAGACAGTTGAGCTGGGCTCTGGGTTGGGATGTGACATTTGACAGTCT 1080
DB 1348 TTCTGCTGCAGACAGTTGAGCTGGGCTCTGGGTTGGGATGTGACATTTGACAGTCT 1407
QY 1031 GCCCATGTACAAAGTGAACCTATACAGTAAACAGTGGCAGTGAAGGACACCTGCC 1140
DB 1408 GCCCATGTACAAAGTGAACCTATACAGTAAACAGTGGCAGTGAAGGACACCTGCC 1467
QY 1141 AAAAATTCACCTGCAGCAGCTTCCAGAGAGAGCTGCATCTGACGGGCTTCGATCGCA 1200
DB 1468 AAAAATTCACCTGCAGCAGCTTCCAGAGAGAGCTGCATCTGACGGGCTTCGATCGCA 1527
QY 1201 TTACTGTATAGAGAAATATCCAACTGATGATGTAATTTGCTGCTGACGAT 1260
DB 1528 TTACTGTATAGAGAAATATCCAACTGATGATGTAATTTGCTGCTGACGAT 1587
QY 1261 GGGGAAGCAACACTATATAGGGGGCTTTAAGCAGGTCAAAGAGTGGCATATC 1320
DB 1588 GGGGAAGCAACACTATATAGGGGGCTTTAAGCAGGTCAAAGAGTGGCATATC 1647
QY 1321 CACACAGTCTGCTTGGGCTCTGACCTCAGCTCAGAACTAGAGAGCTGTCCAAATGACA 1380
DB 1648 CACACAGTCTGCTTGGGCTCTGACCTCAGCTCAGAACTAGAGAGCTGTCCAAATGACA 1707
QY 1381 GAGGTTTACAGACATATGCTTCAGATCAAGTTCAGAAATGAGCTCTATGATGCTTT 1440
DB 1708 GAGGTTTACAGACATATGCTTCAGATCAAGTTCAGAAATGAGCTCTATGATGCTTT 1767
QY 1441 GGGGCTCTTATCAGAAATGAGAGCTCTCTCAGGCTCATATCAGCTGAGAGTATG 1500
DB 1768 GGGGCTCTTATCAGAAATGAGAGCTCTCTCAGGCTCATATCAGCTGAGAGTATG 1827
QY 1501 GGATTTAACCTCCAGAACAGCCAGTGATGATGAGACAGTATGAGACAGCAGCTG 1560
DB 1828 GGATTTAACCTCCAGAACAGCCAGTGATGATGAGACAGTATGAGACAGCAGCTG 1887
QY 1561 GGAAGGACACTTGTCTTATACCTGAGCAACAGCAGCTCCCAATCTCTCTG 1620
DB 1888 GGAAGGACACTTGTCTTATACCTGAGCAACAGCAGCTCCCAATCTCTCTG 1947
QY 1621 GATCCAGTGGACAGAAAGTGTGCTTGTAGTGAACAAACACCAAAATGGCTTAC 1680
DB 1948 GATCCAGTGGACAGAAAGTGTGCTTGTAGTGAACAAACACCAAAATGGCTTAC 2007
QY 1681 CTCCAATCCAGGCAATGCTAAGGTTGGCACTTGGAATACAGTCTCAAGCAAGCTCA 2067
DB 2008 CTCCAATCCAGGCAATGCTAAGGTTGGCACTTGGAATACAGTCTCAAGCAAGCTCA 2127
QY 1741 CAATCTGACCTGACAGTGTACAGTCCGCTGGTCCATATGCTCCCTCCCAATTCA 1800
DB 2068 CAATCTGACCTGACAGTGTACAGTCCGCTGGTCCATATGCTCCCTCCCAATTCA 2127
QY 1801 GTGACTTCCAAACGAAAGAGCAGCAGCAAAATCCCGACCTCTGTGATTATGCA 1860
DB 2128 GTGACTTCCAAACGAAAGAGCAGCAGCAAAATCCCGACCTCTGTGATTATGCA 2187
QY 1861 AATATTTGCCAAGAGCTCCCAATTTCCAGGCGACAGTGTACAGCCCTATGAAATCA 2247
DB 2188 AATATTTGCCAAGAGCTCCCAATTTCCAGGCGACAGTGTACAGCCCTATGAAATCA 2307
QY 1921 GTGATGAAAAAGAGTATGCTGAGTATGAGTATGAGAGGAGGATGATGCTACT 1980
DB 2248 GTGATGAAAAAGAGTATGCTGAGTATGAGTATGAGAGGAGGATGATGCTACT 2307
QY 1981 AAGGATGAGAGTGTCTACTCAAGGATTTTCACTATGACAGCAATGATGATACAGT 2040
DB 2308 AAGGATGAGAGTGTCTACTCAAGGATTTTCACTATGACAGCAATGATGATACAGT 2367


```

QY 2041 GTAAAGTCCGGCTCTGGGAGAGATTAAACGACCCAGAGAGATGATACCCAGAG 2100
DB 2368 TTTAAAGTCCGGCTCTGGGAGAGATTAAACGACCCAGAGAGATGATACCCAGAG 2427
QY 2101 AGGAGGACCTGTACATCTGCTGGATGTAATGATGTAATGATGATGATGATGAT 2160
DB 2428 AGGAGGACCTGTACATCTGCTGGATGTAATGATGTAATGATGATGATGATGAT 2487
QY 2161 AGACGTAATTAATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 2220
DB 2488 AGACGTAATTAATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 2547
QY 2221 TCGGAGGCTCATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 2280
DB 2548 TCGGAGGCTCATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 2607
QY 2281 CCGGAGGCTCATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 2340
DB 2608 CCGGAGGCTCATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 2667
QY 2341 TGGAGAGCTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
DB 2668 TGGAGAGCTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2727
QY 2401 AGTACAGTATTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2460
DB 2728 AGTACAGTATTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2787
QY 2461 GCTCTATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
DB 2788 GCTCTATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2847
QY 2521 ACTTTGAAAATGGCAGATCTTTTCAATTCATTCATTCATTCATTCATTCATTCAT 2580
DB 2848 ACTTTGAAAATGGCAGATCTTTTCAATTCATTCATTCATTCATTCATTCATTCAT 2907
QY 2581 AAATCAGAAATATCAACATTCGACAGATCTTTTCAATTCATTCATTCATTCAT 2640
DB 2908 AAATCAGAAATATCAACATTCGACAGATCTTTTCAATTCATTCATTCATTCAT 2967
QY 2641 GAGACAGCTAGTCTGATGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
DB 2968 GAGACAGCTAGTCTGATGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3027
QY 2701 ATTCTGCGATTCACATTTTAAATTAATGATGATGATGATGATGATGATGATGAT 2760
DB 3028 ATTCTGCGATTCACATTTTAAATTAATGATGATGATGATGATGATGATGATGAT 3087
QY 2761 ATAGCTAGGCTGAATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 2820
DB 3088 ATAGCTAGGCTGAATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 3146
QY 2821 TTAATAAATTTTCTAATAATGATGATGATGATGATGATGATGATGATGATGAT 2880
DB 3147 TTAATAAATTTTCTAATAATGATGATGATGATGATGATGATGATGATGATGAT 3206
QY 2881 ATAGTACATTTTACTAATAATGATGATGATGATGATGATGATGATGATGATGAT 2940
DB 3207 ATAGTACATTTTACTAATAATGATGATGATGATGATGATGATGATGATGATGAT 3266
QY 2941 TTCTGTAGGGGGCGATTAATAATGCTAAACATGCGGTA 2983
DB 3267 TTCTGTAGGGGGCGATTAATAATGCTAAACATGCGGTA 3309

```

```

RESULT 2
AF039400 3311 bp mRNA PRI 14-DEC-1998
LOCUS Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA,
DEFINITION complete cds.
ACCESSION AF039400
VERSION AF039400.1 GI:4009457
KEYWORDS

```

```

SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3311)
AUTHORS Gruber,A.D., Eble,R.C., Ji,H.L., Schreur,K.D., Fuller,C.M. and
Pauli,B.U.
TITLE Genomic cloning, molecular characterization, and functional
analysis of human CLCA1, the first human member of the family of
Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 9904/526
REFERENCE 2 (bases 1 to 3311)
AUTHORS Gruber,A.D., Eble,R.C. and Pauli,B.U.
TITLE Direct Submision
JOURNAL Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
Cornell University, Ithaca, NY 14853, USA
FEATURES
source
1. 3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/tissue_type="small intestine"
1. 3311
/gene="hCLCA1"
352..3096
/feature="hCLCA1"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPKRSVPLILHLLEGALSLSLIQLNNNGEIVAVDNPV
EDETLIQIKDNTQASLYLPERGKRTFKVAVILIPETWTKADYVPRKLETTANA
DIVIAESTPGNDPEYEQMGNGEGERIHLTPDIACKLAEGPQKAFVHEWAH
LRNGVDEYNDEKPEYLSNGRIQAVCSGIGTNNVKKCOGSCYTKCTENKVG
YKGECEFLVQSROTERKASIMFQOHVDSIYEPCEONHNEAPKOKCNLSSTWEVI
RSEDEKKTTPMTOPNPTELSLQIGORIVCLDKSGSMATGNLNLNNGOLEFL
LOTVELSGVNTFDSAAHVOSELIQINSRDRILAKRLPAASGRTSICGLPSA
FYIRKRYTDSGEIYLLDGEENTISGCFNEPKOSGALIHVALPASAQELBELSK
MTGGLQTVASDOVONNGILDAFSAISGNAVASQISQLESGLTVLQNSQMNNGTVI
DSVGCADTLEFLTWTPQPOLILMDPSGKOGQFVMDKTKAAYLQIPILAKGTWKY
SLOASQOTLTVTSRASNALEPIPIVTSKTKNDISKPPSPVAVYANIRQASPLIRA
SVYALIESVNGKIVLELDNNGADATDDGYSHYFTFYDNGRYSVKVALGGVN
AARRYIPQOSGLYIPGMIENDEIQMNPREINDVOYHOVCPSSSGSPVAS
DVPNAPIPLPFPFGQITDLKAEIHGSLNLNLTATAGDDYDHTAKYIIRISTSLD
LRDNESLQVNTALILPKEANSEVFLKPERNTEENGDTLTAIOADVKNLSEI
SNARVSLRIPQTPPTPSPDPTSAFCPNHINSTIPGIIHLIKMKWIGELQLSIA
"
BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN
Query Match 99.4%; Score 2966.2; DB 9; Length 3311;
Best local Similarity 99.9%; Pred. No. 0;
Matches 2979; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```


0Y 241 TATTTCAAAATGTTGCCATTTTGATTCCTGGAACATGGAAGACAAGACGCTGACTATGTG 300
| | | | |
Db 568 TATTTCAAAATGTTGCCATTTTGATTCCTGGAACATGGAAGACAAGACGCTGACTATGTG 627
0Y 301 AGACCAAAACCTTGAGACCTACAAAAATGCTGATGTTCTGGTGTGAGTCTACTCTCTCA 360
| | | | |
Db 628 AGACCAAAACCTTGAGACCTACAAAAATGCTGATGTTCTGGTGTGAGTCTACTCTCTCA 687
0Y 361 GGTAAATGATGAACCTTACACTGAGAGATGGGCAACTGTGAGAGAGAGGTTGAAAGATC 420
| | | | |
Db 688 GGTAAATGATGAACCTTACACTGAGAGATGGGCAACTGTGAGAGAGAGGTTGAAAGATC 747
0Y 421 CACCTCACTCTGATTTGATTCGAGAAAAAGTTAGCTGAATATGACCAAGATGAG 480
| | | | |
Db 748 CACCTCACTCTGATTTGATTCGAGAAAAAGTTAGCTGAATATGACCAAGATGAG 807
0Y 481 GCATTTGTCATGATGAGGCTCATCTACGATGGGAGTATTTGACGAGTACATATGAT 540
| | | | |
Db 808 GCATTTGTCATGATGAGGCTCATCTACGATGGGAGTATTTGACGAGTACATATGAT 867
0Y 541 GAGAAATTTCTACTATTCATGAGAAATACAGCATAGATGTTTCAGCAGCTATTTCT 600
| | | | |
Db 868 GAGAAATTTCTACTATTCATGAGAAATACAGCATAGATGTTTCAGCAGCTATTTCT 927
0Y 601 GGTACAAATGTGTAGAGAGTGTGAGGAGGAGGAGCTGTACACCAAAAGATGACATTC 660
| | | | |
Db 928 GGTACAAATGTGTAGAGAGTGTGAGGAGGAGGAGCTGTACACCAAAAGATGACATTC 987
0Y 661 AATAAGTAAACAGACTATGAAAAAGATGTGAGTTGTTCTCAATCCCGCAGAGG 720
| | | | |
Db 988 AATAAGTAAACAGACTATGAAAAAGATGTGAGTTGTTCTCAATCCCGCAGAGG 1047
0Y 721 GGAAGGCTTCTATATGTTTGACACAATGTTGATTCATGTAATTCGTACAGA 780
| | | | |
Db 1048 GGAAGGCTTCTATATGTTTGACACAATGTTGATTCATGTAATTCGTACAGA 1107
0Y 781 CAAAACCAACAAAGAAAGAGCTCCAAACAGCAAAATCAAAATGCAATTCGGAAGCACA 840
| | | | |
Db 1108 CAAAACCAACAAAGAAAGAGCTCCAAACAGCAAAATCAAAATGCAATTCGGAAGCACA 1167
0Y 841 TGGGAAGTATCCGTGATTCGTGAGCACTTTAAGAAAAACCACTCTATGCAACACAGCCA 900
| | | | |
Db 1168 TGGGAAGTATCCGTGATTCGTGAGCACTTTAAGAAAAACCACTCTATGCAACACAGCCA 1227
0Y 901 CCAATTCACACCTCTCATTCATTCGTGAGATGTCGACAAAGATGTTGTTAGCTTGCAC 960
| | | | |
Db 1228 CCAATTCACACCTCTCATTCATTCGTGAGATGTCGACAAAGATGTTGTTAGCTTGCAC 1287
0Y 961 AATATCTGAAGATGGCGAGCTGTAAACGCTCAATCGACTGAATCAGCAGGCCAGCTT 1020
| | | | |
Db 1288 AATATCTGAAGATGGCGAGCTGTAAACGCTCAATCGACTGAATCAGCAGGCCAGCTT 1347
0Y 1021 TTCCTGCTCAGACAGTGTAGCTGGGGTCTGGGTTGGGATGTGCAATTTGACATGCT 1080
| | | | |
Db 1348 TTCCTGCTCAGACAGTGTAGCTGGGGTCTGGGTTGGGATGTGCAATTTGACATGCT 1407
0Y 1081 GCCCATGTACAAAGTAACTCATACAGATTAACAGTGGCAGTGAACAGGAGCACTCGCC 1140
| | | | |
Db 1408 GCCCATGTACAAAGTAACTCATACAGATTAACAGTGGCAGTGAACAGGAGCACTCGCC 1467
0Y 1141 AAAAGATTAACCTGACAGCTTCAGAGGAGACGCTCATCTGACGCGGCTTGATGGCA 1200
| | | | |
Db 1468 AAAAGATTAACCTGACAGCTTCAGAGGAGACGCTCATCTGACGCGGCTTGATGGCA 1527
0Y 1201 TTTTACGTATTTAGGAAGAAATTCACAGTGAATCTGAAATTTGCTGCTGACGGAT 1260
| | | | |
Db 1528 TTTTACGTATTTAGGAAGAAATTCACAGTGAATCTGAAATTTGCTGCTGACGGAT 1587
0Y 1261 GGGGAGAGACAACATAAAGTGGGCTTTAAGGAGTCAACAAAGTGGTGGCATCATC 1320
| | | | |
Db 1588 GGGGAGAGACAACATAAAGTGGGCTTTAAGGAGTCAACAAAGTGGTGGCATCATC 1647

0Y 1321 CACACAGTCTGTTGGGCCCCCTCTGACAGCTCAAGACATAGAGAGCTGTCCAAAATGACA 1380
| | | | |
Db 1648 CACACAGTCTGTTGGGCCCCCTCTGACAGCTCAAGACATAGAGAGCTGTCCAAAATGACA 1707
0Y 1381 GGAGTTTACAGACATATGCTTCAGATCAAGTCAAGAAATGAGCTCATATTTGCTTTT 1440
| | | | |
Db 1708 GGAGTTTACAGACATATGCTTCAGATCAAGTCAAGAAATGAGCTCATATTTGCTTTT 1767
0Y 1441 GGGGCCCCCTTCATCAGGAAGAAATGAGAGCTGTCTCAGAGGCTTCATCCAGCTTGAGAGTAA 1500
| | | | |
Db 1768 GGGGCCCCCTTCATCAGGAAGAAATGAGAGCTGTCTCAGAGGCTTCATCCAGCTTGAGAGTAA 1827
0Y 1501 GGATTAACCTCCAGAACAGCAGCTGATGATGAGCAACAGTATGATGAGACAGCAGCTG 1560
| | | | |
Db 1828 GGATTAACCTCCAGAACAGCAGCTGATGATGAGCAACAGTATGATGAGACAGCAGCTG 1887
0Y 1561 GGAAGAGACACTTGTGTTCTTATACACTGAGCAACGACGCTCCCAATCTCTCTGG 1620
| | | | |
Db 1888 GGAAGAGACACTTGTGTTCTTATACACTGAGCAACGACGCTCCCAATCTCTCTGG 1947
0Y 1621 GATCCACAGTGAACAGAACAGAGTGGGCTTTGTAGTGAACAAAACACCAAAATGGCTTAC 1680
| | | | |
Db 1948 GATCCACAGTGAACAGAACAGAGTGGGCTTTGTAGTGAACAAAACACCAAAATGGCTTAC 2007
0Y 1631 CTCCAATCCAGAGCATTTGCTTAAGTTGGCACTTGAAATACAGTCTGCAAGACAGCTCA 1740
| | | | |
Db 2008 CTCCAATCCAGAGCATTTGCTTAAGTTGGCACTTGAAATACAGTCTGCAAGACAGCTCA 2067
0Y 1741 CAACCTTGACCTTGACTGTACAGTCCGCTGCTCCATAGCTACCTCCCTCCCAATTACA 1800
| | | | |
Db 2068 CAACCTTGACCTTGACTGTACAGTCCGCTGCTCCATAGCTACCTCCCTCCCAATTACA 2127
0Y 1801 GTGACTTCCAAAAGCAAGAACAGACACAGCAAAATTCGCCAGCCCTGTGATTTATGCA 1860
| | | | |
Db 2128 GTGACTTCCAAAAGCAAGAACAGACACAGCAAAATTCGCCAGCCCTGTGATTTATGCA 2187
0Y 1861 AATATTCGCAAGAGAGCTCCCAATTTCTCAGGGCAGTATGACAGCCCTATGTAATCA 1920
| | | | |
Db 2188 AATATTCGCAAGAGAGCTCCCAATTTCTCAGGGCAGTATGACAGCCCTATGTAATCA 2247
0Y 1921 GTGATGGAAGAAACAGTATACCTTGGAATCTAGTATGAGAGAGTGTGATGCTACT 1980
| | | | |
Db 2248 GTGATGGAAGAAACAGTATACCTTGGAATCTAGTATGAGAGAGTGTGATGCTACT 2307
0Y 1981 AAGGATGACGGTGTCTACTCAAGTATTTCACTATGACAGATGATGATACAGT 2040
| | | | |
Db 2308 AAGGATGACGGTGTCTACTCAAGTATTTCACTATGACAGATGATGATACAGT 2367
0Y 2041 GTAAAAGTCCGGGCTCTGGGAGAGTAAACGAGCAGCAGAGAGTATACCCAGAG 2100
| | | | |
Db 2368 GTAAAAGTCCGGGCTCTGGGAGAGTAAACGAGCAGCAGAGAGTATACCCAGAG 2427
0Y 2101 AGTGAAGCTGTACATACCTGGCTGATGATGAGTGAATGAATGAATGCAATCCACA 2160
| | | | |
Db 2428 AGTGAAGCTGTACATACCTGGCTGATGATGAGTGAATGAATGAATGCAATCCACA 2487
0Y 2161 AGACCTGAATTAATPAAGATGATGTTCAACAAAGAAATGTTTACGAGAAATCC 2220
| | | | |
Db 2488 AGACCTGAATTAATPAAGATGATGTTCAACAAAGAAATGTTTACGAGAAATCC 2547
0Y 2221 TGGGAGAGCTCATTTGTGGCTCTGATGTCACCAATGCTCCCATCTGATCTTCCA 2280
| | | | |
Db 2548 TGGGAGAGCTCATTTGTGGCTCTGATGTCACCAATGCTCCCATCTGATCTTCCA 2607
0Y 2281 CCTGGCAAAATACAGACCTGGAAGCGGAAATTCACGGGGGAGCTCATTAATCTGACT 2340
| | | | |
Db 2608 CCTGGCAAAATACAGACCTGGAAGCGGAAATTCACGGGGGAGCTCATTAATCTGACT 2667
0Y 2341 TGGACAGTCTCTGGGGATGATATGACATGGAACAGCTCAACAATATATCTTTGATA 2400
| | | | |
Db 2668 TGGACAGTCTCTGGGGATGATATGACATGGAACAGCTCAACAATATATCTTTGATA 2727
0Y 2401 AGTACAAATTTCTGATCTCAGAGCAAGTTCATGAAATCTCTTCAAGTGAATCTACT 2460
| | | | |

Db 2728 ACTACAAAGTATCTTGATCTCAGACAGCAAGTCAATGATCTTCAAGGAACTACTACT 2787
 Oy 2461 GCTTCATCCCAAGAGAGCAACTCTGAGAGAGCTTTTGTAAACAGAAACATT 2520
 Db 2788 GCTTCATCCCAAGAGAGCAACTCTGAGAGAGCTTTTGTAAACAGAAACATT 2520
 Oy 2521 ACTTTGAAAATGCGACAGATCTTTTCAATTCAGTCTGATTAAGTGCATCTG 2580
 Db 2848 ACTTTGAAAATGCGACAGATCTTTTCAATTCAGTCTGATTAAGTGCATCTG 2580
 Oy 2581 AATTCGAAAATTCGACAGATCTTTTCAATTCAGTCTGATTAAGTGCATCTG 2580
 Db 2908 AATTCGAAAATTCGACAGATCTTTTCAATTCAGTCTGATTAAGTGCATCTG 2580
 Oy 2641 GAGACACCTAGTCTGATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
 Db 2968 GAGACACCTAGTCTGATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
 Oy 2701 ATTCTGCGATTCACATTTTAAATTAATGAGAGAGTGAAGAGAGTGCAGTCTG 2760
 Db 3028 ATTCTGCGATTCACATTTTAAATTAATGAGAGAGTGAAGAGAGTGCAGTCTG 2760
 Oy 2761 ATAGCTAGGCGTGAATTTTGTGAGATTAATTAATTAATTAATTAATTAATTA 2820
 Db 3088 ATAGCTAGGCGTGAATTTTGTGAGATTAATTAATTAATTAATTAATTAATTA 2820
 Oy 2821 TATATAAATTTTCTAAATGATTTTGTAGACTCTGCTGCTGCTGCTGCTGCTGCT 2880
 Db 3147 TATATAAATTTTCTAAATGATTTTGTAGACTCTGCTGCTGCTGCTGCTGCTGCT 2880
 Oy 2881 ATAGTACATTTTCTAAATGATTTTGTAGACTCTGCTGCTGCTGCTGCTGCTGCT 2940
 Db 3207 ATAGTACATTTTCTAAATGATTTTGTAGACTCTGCTGCTGCTGCTGCTGCTGCT 2940
 Oy 2941 TTCCCTGAGGGGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2983
 Db 3267 TTCCCTGAGGGGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2983

RESULT 3
 AF127036 2826 bp mRNA PRI 10-AUG-1999
 LOCUS AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CACCI)
 DEFINITION Homo sapiens calcium-activated chloride channel protein 1 (CACCI)
 ACCESSION AF127036
 VERSION AF127036.1 GI:4585468
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2826)
 AUTHORS Agnel, M., Vernat, T. and Culouscou, J.-M.
 TITLE Identification of three novel members of the calcium-dependent
 chloride channel (CACCI) family predominantly expressed in the
 digestive tract and trachea
 JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
 MEDLINE 99364503
 PUBMED 10437792
 REFERENCE 2 (bases 1 to 2826)
 AUTHORS Agnel, M. and Culouscou, J.-M.
 TITLE Direct Submission
 JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des
 Carrieres, Rueil-Malmaison 92500, France
 FEATURES
 source
 1. .2826
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="small intestine; colon"
 1. .2826
 /gene="CACCI"
 CDS
 5. .2749

BASE COUNT 875 a 623 c 632 g 696 t
 ORIGIN
 Query Match 94.1%; Score 2806.8; DB 9; Length 2826;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2811; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 /gene="CACCI"
 /note="Dovine epithelial chloride channel homolog"
 /product="calcium-activated chloride channel protein 1"
 /protein_id="AA025487.1"
 /db_xref="GI:4585469"
 /translation="MGPKRSVFILHLLEGALSNSLIOLNNNGYEGIVAVIDPNVP
 EDELTLQIKDVTQASLYLEFATKREFKNVALLIPETWTKADYVPKLETKNA
 DVLVASETPPGDEPDEBMGNGKGEIHLTPGIGKKLAEPGGRVHMHM
 LEMVPEDEVNDEKPYLSRGIOARVSGGATGIVYVKKCGSGCYTRCTRTGL
 YKGEFVLQISQTERKASIMFQHDYSIVERTEDONHKEARNKOKNLSIEVI
 RQSEDFKTPPTPTOPNPFTSLDVGIVLVLDKSGSMATGNRLNUNAGOLF
 LQTELVGSVNGVTPDSAAHVOSELIOINSGRDPTLAKRLPAAASGTSISGLRSA
 FTVIRKKYPTDSEIVLLDGEDNTISGCFNORSOGALIHVALGPSAOEELISK
 MTGKLOTASDOVONNGILDAFGALSNGGAVSORSIOLESKGLLONSOMNGTVI
 DSTVGKTLFTTMTTOPPOLILMPSGKOGFVVDKRTKAYLQIPGIAKVTWKY
 SIQASSQTLITLVTSBASNATIPITVTSKTKNDPSKPSPLVYANIKQASPLIRA
 SVTALIESNNGTVLELDNAGADATDKDVSRYFTYPTNGRYSKVVALAGVN
 AARRVYIPQSGALYIPWIMENDEIOMNPREDINKDVOHGVFSRTSSGSEVAS
 DVNPADIPDLPPGOITDLKABIHGGSFLNLTWPAAGDDYDHGTAKHYIRISTILD
 LRDKNESLOVNTFTALIPREANSEVLEKPEINTEPENGTDLEIALIOAVDKYLSKI
 SNIRVSLFTIPQTPPEPSPBETSPAPCNHININSTIRPQIHILKIMKIGELQSLIA
 "
 1. 21 ACCAATGGGGCCATTAAAGAGTCTGTGTCATCTGATTTCTTACCTTCTAGAGGGC 80
 2. 1 AGCAATGGGGCCATTAAAGAGTCTGTGTCATCTGATTTCTTACCTTCTAGAGGGC 60
 3. 1 GGCATCTCTGATCTGTGTTGAAGCTACAGGAAGGCAATTTATTTCAAAATGTTCCAT 260
 4. 1 GGCATCTCTGATCTGTGTTGAAGCTACAGGAAGGCAATTTATTTCAAAATGTTCCAT 240
 5. 1 TTGATTTCTGAAACATGGAAGCAAAAGGCTGACTATGTGAGACCAAACTTGAGACTA 300
 6. 1 CAAAATGCGATGTTCTGTTGCTGAGTCTGCTGAGTCTGCTGCAAGTAATGATGAACCTGAC 380
 7. 1 CAAAATGCGATGTTCTGTTGCTGAGTCTGCTGAGTCTGCTGCAAGTAATGATGAACCTGAC 360
 8. 1 TGACGAGATGGGCACTGTGAGAGAGGTTGAAGAGTACACCTGCTGATTTGAT 440
 9. 1 TGACGAGATGGGCACTGTGAGAGAGGTTGAAGAGTACACCTGCTGATTTGAT 420
 10. 1 TGACGAGAGAGAGTGTAGCTGAATATGACACCAAGGTAGGGCATTTGTCCATGAGTGGC 500
 11. 1 TGACGAGAGAGAGTGTAGCTGAATATGACACCAAGGTAGGGCATTTGTCCATGAGTGGC 500
 12. 1 TGACGAGAGAGAGTGTAGCTGAATATGACACCAAGGTAGGGCATTTGTCCATGAGTGGC 480
 13. 1 TCATCTACGATGGGAGTATTTGACGAGTCAATTAATGATGAAGAAATTTGATTTACCA 560
 14. 1 TCATCTACGATGGGAGTATTTGACGAGTCAATTAATGATGAAGAAATTTGATTTACCA 540
 15. 1 TGCAAGAAATACAGCAGTAAAGATGTTACAGAGTATTTACTGTACAAATGTTACTAAGAA 620
 16. 1 TGCAAGAAATACAGCAGTAAAGATGTTACAGAGTATTTACTGTGTACAAATGTTACTAAGAA 600
 17. 1 GTGTACAGAGGCGAGCTGTTACACCAAGATGACATTTCAATTAAGTAAAGTACGAGACTTA 680

D	601	GTGTCAGGAGGACGACGTGTACACCCAAAGATGACATTTCAATTAAGATTACAGACTCTA	660
O	681	TGAAAAAGATGTGAGCTTTGTTCTTCCAAATCCCGCCACAGGAGAGGCTCTATTAATGTT	740
D	661	TGAAAAAGATGTGAGTTTGTCTTCCAAATCCCGCCACAGGAGAGGCTCTATTAATGTT	720
O	741	TGCACCAACATGTTGATCTCTATGTTGTAATTCGTATGACAGAAACCAACACAAACAGC	800
D	721	TGCACCAACATGTTGATCTCTATGTTGTAATTCGTATGACAGAAACCAACACAAACAGC	780
O	801	TCCAAACAGCAAAATCAAAAAATGCAATCTCCGAAGACATGGGAAGTGAATCCGTATTC	860
D	781	TCCAAACAGCAAAATCAAAAAATGCAATCTCCGAAGACATGGGAAGTGAATCCGTATTC	840
O	861	TGAGAGCTTTAAGAAAACCACTCTATGACAAACAGCACGCCAACCAATCCCACTTCATAT	920
D	841	TGAGAGCTTTAAGAAAACCACTCTATGACAAACAGCACGCCAACCAATCCCACTTCATAT	900
O	921	GCTGCACATTTGACAAAGAATGTGTGTTTGTGCTTGACAAATCTGGAAGCATGGCGAC	980
D	901	GCTGCACATTTGACAAAGAATGTGTGTTTGTGCTTGACAAATCTGGAAGCATGGCGAC	960
O	981	TGGTAAACCGCTCAATGCACTGATATCAAGACAGCCAGCTTTTCTGCTGCAAGCACTTGA	1040
D	961	TGGTAAACCGCTCAATGCACTGATATCAAGACAGCCAGCTTTTCTGCTGCAAGCACTTGA	1020
O	1041	GCTGGGTCTCTGGGTTGGGATGGTGATCAATTTACAGTCTGCCCCATGTACAAAGTGAAT	1100
D	1021	GCTGGGTCTCTGGGTTGGGATGGTGATCAATTTACAGTCTGCCCCATGTACAAAGTGAAT	1080
O	1101	CATACAGATTAAGACAGTGGCGAGTGCACAGGGACACACTGCCCAAAAGATTACTGCAAGC	1160
D	1081	CATACAGATTAAGACAGTGGCGAGTGCACAGGGACACACTGCCCAAAAGATTACTGCAAGC	1140
O	1161	TTTCAGGAGGACGTCATCTGCAGCGGGGTTGTGATCGGCATTTACTGTGATTAGAGAA	1220
D	1141	TTTCAGGAGGACGTCATCTGCAGCGGGGTTGTGATCGGCATTTACTGTGATTAGAGAA	1200
O	1221	ATATCCAACTGATGATCTGAAATTTGTGCTGTGACGGATGGGGAAAGCAACACTATTAAG	1280
D	1201	ATATCCAACTGATGATCTGAAATTTGTGCTGTGACGGATGGGGAAAGCAACACTATTAAG	1260
O	1281	TGGGTGCTTTAAGCAAGGTCAAAACAAAGTGGTGCATTCACACAGTGCCTTTGGGGCC	1340
D	1261	TGGGTGCTTTAAGCAAGGTCAAAACAAAGTGGTGCATTCACACAGTGCCTTTGGGGCC	1320
O	1341	CTCTGCAGCTCAAGAACTAAGAGAGCTGTCCAAATGACAGAGGTTTACAGACATATGC	1400
D	1321	CTCTGCAGCTCAAGAACTAAGAGAGCTGTCCAAATGACAGAGGTTTACAGACATATGC	1380
O	1401	TTTCAGATCAAGTTCAAGAAACAATGGCTCATTTGATGCTTTTGGGGCCCTTATACAGAAA	1460
D	1381	TTTCAGATCAAGTTCAAGAAACAATGGCTCATTTGATGCTTTTGGGGCCCTTATACAGAAA	1440
O	1461	TGGAGCTCTCTCAGAGCTTCCATCCAGCTTGAGATTAAGGATTATACCTCCAGAACAG	1520
D	1441	TGGAGCTCTCTCAGAGCTTCCATCCAGCTTGAGATTAAGGATTATACCTCCAGAACAG	1500
O	1521	CCAGTGGATGAATGGCAACAGTGAATCGTGGACAGCACCGTGGAAAGACACTTGTGTTCT	1580
D	1501	CCAGTGGATGAATGGCAACAGTGAATCGTGGACAGCACCGTGGAAAGACACTTGTGTTCT	1560
O	1581	TTTTCACCTGGCAACGACAGCTCCCAAAATCTCTCTGGAATCCAGTGCAGAGAACGA	1640
D	1561	TTTTCACCTGGCAACGACAGCTCCCAAAATCTCTCTGGAATCCAGTGCAGAGAACGA	1620
O	1641	AGGTGGCTTTGATGACAAAAACACCAAAATGGCTTACCTCCAAATCCAGCATTCG	1700
D	1621	AGGTGGCTTTGATGACAAAAACACCAAAATGGCTTACCTCCAAATCCAGCATTCG	1680
O	1701	TAAAGTTGGACATTGGAAATTAAGTGTGCAGAGCAAGCACTCAAAACCTTGACCTGACTGT	1760
D	1681	TAAAGTTGGACATTGGAAATTAAGTGTGCAGAGCAAGCACTTGACCTGACTGACTGT	1740

[illegible]

RESULT 4
AK024970 2022 bp mRNA PRI 29-SEP-2000
LOCUS Homo sapiens cDNA: FLJ21317 fls, clone COL02275, highly similar to
DEFINITION AF127036 Homo sapiens calcium-activated chloride channel protein 1
(CaCC1) mRNA.
ACCESSION AK024970
VERSION AK024970.1 GI:10437397
KEYWORDS oligo cloning: fls (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone lib:COL clone:COL02275.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (sites)
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
MEDO human cDNA sequencing project
Unpublished (2000)
2 (bases 1 to 2022)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Fax:81-3-5449-5416)
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo,
Laboratory of Genome Structure Analysis, Human Genome Center;
Shitokone-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
FAX:81-3-5449-5416)
COMMENT MEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing; Research Association for Biotechnology: cDNA library
construction; 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source Location/Qualifiers
1..2022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL02275"
/clone_lib="COL"
/tissue_type="colon"
/note="cloning vector pME18SFL3"
misc_feature 1..2022
/note="highly similar to AF127036 Homo sapiens
calcium-activated chloride channel protein 1 (CaCC1) mRNA"
BASE COUNT 612 a 472 c 453 g 485 t
ORIGIN
Query Match 67.0%; Score 1999.4; DB 9; Length 2022;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2006; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
OY 812 AAAATCAAAAATGCAATCTCCGAAGACATGGAAGTATCGCTGAGAGACTTTA 871
DB 1 AAAATCAAAAATGCAATCTCCGAAGACATGGAAGTATCGCTGAGAGACTTTA 60
OY 872 AGAAAACACCTCTATGACAAACACACACCAATCCACCTTCTCATTTGCTCAGATTG 931
DB 61 AGAAAACACCTCTATGACAAACACACCAATCCACCTTCTCATTTGCTCAGATTG 120
OY 932 GCAAAAGATTTGTTGTTAGTCTTGACAAATCTGGAACATCGGACATGTAACCGCC 991
DB 121 GCAAAAGATTTGTTGTTAGTCTTGACAAATCTGGAACATCGGACATGTAACCGCC 180
OY 992 TCAATGCAATGCAATGCAAGACGACGCTTTTCTGCTGACAGAGTTGAGCTGGGCTCT 1051
DB 181 TCAATGCAATGCAATGCAAGACGACGCTTTTCTGCTGACAGAGTTGAGCTGGGCTCT 240
OY 1052 GGGTTGGATGTTGACATTTGACAGTGTGCCCCATGTACAAAGTGAATCATACAGATA 1111

|||||
db 111 GGGTTGGATGTTGACATTTGACAGTGTGCCCCATGTACAAAGTGAATCATACAGATA 300
OY 112 ACAGTGGAGTGAACAGGAGACACACTCGCCAAAAGATTACTGAGACAGCTTCAGAGGGA 1171
DB 301 ACAGTGGAGTGAACAGGAGACACACTCGCCAAAAGATTACTGAGACAGCTTCAGAGGGA 360
OY 1172 CGTCAATGTGACAGCGGCTTCGATCGGCAATTTACTGATTAGAGAAATATCAACTG 1231
DB 161 CGTCAATGTGACAGCGGCTTCGATCGGCAATTTACTGATTAGAGAAATATCAACTG 420
OY 1232 ATGATCTGAATTTGCTGCTGCTGACAGGATGGGAAGACAACTATTAAGTGGTCTTTA 1291
DB 421 ATGATCTGAATTTGCTGCTGCTGACAGGATGGGAAGACAACTATTAAGTGGTCTTTA 480
OY 1292 ACAGGTCAAACAAAGTGGGTCATCATCCACACAGTCGTTGGGCCCCCTGCGACCTC 1351
DB 481 ACAGGTCAAACAAAGTGGGTCATCATCCACACAGTCGTTGGGCCCCCTGCGACCTC 540
OY 1352 AGAAGTGAAGAGAGCTGTCCAAATGACAGAGGTTTACAGATATGCTTCAGATCAAG 1411
DB 541 AGAAGTGAAGAGAGCTGTCCAAATGACAGAGGTTTACAGATATGCTTCAGATCAAG 600
OY 1412 TTGAGAACTATGCTTCATTTGCTTTGGGCCCTTTTCATGAGAAATGAGCTGCT 1471
DB 601 TTGAGAACTATGCTTCATTTGCTTTGGGCCCTTTTCATGAGAAATGAGCTGCT 660
OY 1472 CTCAGCGCTTCATCCAGCTTGAGAGTAAGGATTTAACCTCCAGAACAGCCAGTGTGTA 1531
DB 651 CTCAGCGCTTCATCCAGCTTGAGAGTAAGGATTTAACCTCCAGAACAGCCAGTGTGTA 720
OY 1532 ATGACACAGATGCTGAGAGACAGCCGTGGAAAGACATTTGTTCTTATACCTGGA 1591
DB 721 ATGACACAGATGCTGAGAGACAGCCGTGGAAAGACATTTGTTCTTATACCTGGA 780
OY 1592 CAACGACGCTCCCAAAATCTTCTGAGATCCAGTCGAGACAGAAAGTGGCTTTG 1651
DB 781 CAACGACGCTCCCAAAATCTTCTGAGATCCAGTCGAGACAGAAAGTGGCTTTG 840
OY 1652 TAGTGGCAAAAACACCAAAATGCGCTCACTCCAAATCCAGGCAATTCGTAAGGTGGCA 1711
DB 841 TAGTGGCAAAAACACCAAAATGCGCTCACTCCAAATCCAGGCAATTCGTAAGGTGGCA 900
OY 1712 CTGGAATACAGTCTGCAAGCAAGCTCACAACTTGACCTGATGCTGACGCTCCGCTG 1771
DB 901 CTGGAATACAGTCTGCAAGCAAGCTCACAACTTGACCTGATGCTGACGCTCCGCTG 960
OY 1772 GGTCAATGCTACCTGCTGCTCAATTAAGTACCTTCCAAAGAACAGACACACGCA 1831
DB 961 GGTCAATGCTACCTGCTGCTCAATTAAGTACCTTCCAAAGAACAGACACACGCA 1020
OY 1832 AATTCGCCAGCCCTGCTGTTTATGCAAAATTTGCGCAAGAGAGCTCCCAATTCGA 1891
DB 1021 AATTCGCCAGCCCTGCTGTTTATGCAAAATTTGCGCAAGAGAGCTCCCAATTCGA 1080
OY 1892 GGGCCAGTGTCAACAGCCCTGATTGAATCACTGATGAGAAAGATTAACCTTGAATAC 1951
DB 1081 GGGCCAGTGTCAACAGCCCTGATTGAATCACTGATGAGAAAGATTAACCTTGAATAC 1140
OY 1952 TGGATTAATGAGAGAGTGTGATGCTAATAAGATGAGGCTGTCTACTCAAGTATTTC 2011
DB 1141 TGGATTAATGAGAGAGTGTGATGCTAATAAGATGAGGCTGTCTACTCAAGTATTTC 1200
OY 2012 CAACCTTAAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2071
DB 1201 CAACCTTAAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
OY 2072 CAGCCAGACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2131
DB 1261 CAGCCAGACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
OY 2132 AGAATGATGAATTAACAATGGAATTCACCAAGACCTGAATTAATTAAGATGATGTTCAAC 2191
|||||

Db	1321	AGATGATGAAATTCATGGAATCCAAAGACTGAAATTAATTAAGATGATGTTCAAC	1380
Qy	2192	ACAAGCAAGTGTGTTTACAGACGAACAATCTCCGGAGGCTATTTTGCGTCTTCAGATGCC	2251
Db	1381	ACAAGCAAGTGTGTTTACAGACGAACAATCTCCGGAGGCTATTTTGCGTCTTCAGATGCC	1440
Qy	2252	CAATTCGTCCTCATTCCTGATCTCTCCACCTGGCCAAATCACCAGACTGAAGCGGAAA	2311
Db	1441	CAATTCGTCCTCATTCCTGATCTCTCCACCTGGCCAAATCACCAGACTGAAGCGGAAA	1500
Qy	2312	TTCAAGGGGGGAGCTCTCATTAATCTGACTTTGACAGAGCTCTGGGGATGATTAATGACCATG	2371
Db	1501	TTCAAGGGGGGAGCTCTCATTAATCTGACTTTGACAGAGCTCTGGGGATGATTAATGACCATG	1560
Qy	2372	GAACAGCTCCACAATATATATCTTGGAAATAGTACAGATTTCTGATCTCAGACACAAT	2431
Db	1561	GAACAGCTCCACAATATATATCTTGGAAATAGTACAGATTTCTGATCTCAGACACAAT	1620
Qy	2432	TCAATGAAATCTCTTCAGATGAAATCTACTGCTCTCATTCCTCCAAAGGAAGCCAACTGTAGG	2491
Db	1621	TCAATGAAATCTCTTCAGATGAAATCTACTGCTCTCATTCCTCCAAAGGAAGCCAACTGTAGG	1680
Qy	2492	AAGCTTTTTTGTTTAAACCAAGAAACATTAATCTTGGAAATAGGACAGATCTTTCATTTG	2551
Db	1681	AAGCTTTTTTGTTTAAACCAAGAAACATTAATCTTGGAAATAGGACAGATCTTTCATTTG	1740
Qy	2552	CTATTCAGAGCTGTTGATTAAGGTGCTGATCTGAATATGAAATATTCACACTTCACAGAGAT	2611
Db	1741	CTATTCAGAGCTGTTGATTAAGGTGCTGATCTGAATATGAAATATTCACACTTCACAGAGAT	1800
Qy	2612	CTTTGTTTATTCCTCCACAGACTCCGCCAGAGACACCTAGTCGATGAAACGTCGTGTC	2671
Db	1801	CTTTGTTTATTCCTCCACAGACTCCGCCAGAGACACCTAGTCGATGAAACGTCGTGTC	1860
Qy	2672	CTTGTCCCTAATATTCATATTCACAAGACACCATTCCTCGGATTCACATTTTAAAAATTATGT	2731
Db	1861	CTTGTCCCTAATATTCATATTCACAAGACACCATTCCTCGGATTCACATTTTAAAAATTATGT	1920
Qy	2732	GGAAGCTGATAGAGAACTGACAGCTGTCATAGCCTAGAGGCTGAATTTTGTACAGATTA	2791
Db	1921	GGAAGCTGATAGAGAACTGACAGCTGTCATAGCCTAGAGGCTGAATTTTGTACAGATTA	1980
Qy	2792	TAAATTAATCATTCATCTTTTTTTGTATTAATTA	2828
Db	1981	TAAATTAATCATTCATCTTTTTTTGTATTAATTA	2017
RESULT	5		
AF095584			
LOCUS	AF095584	3079 bp	mRNA
DEFINITION	Sus scrofa epithelial chloride channel protein (hECC) mRNA,		
ACCESSION	AF095584		
VERSION	AF095584.1	GI:6002645	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			

FEATURES	location/Qualifiers
source	1..3079
	/organism="Sus scrofa"
	/db_xref="taxon:9823"
	/tissue_type="ileal mucosa"
gene	1..3079
	/gene="AECC"
CDS	133..2886
	/gene="AECC"
	/function="Involved in exocrine chloride conductance"
	/codon_start=1
	/product="epithelial chloride channel protein"
	/protein_id="AA00077.1"
	/db_xref="GI:6002646"
	/translation="MGSPRSLSFLTFLVLLHLEGAQNSNLIQLNGNGEGIVIAIDPNVH EDRELINIKDMVTKASPLYLEAEATREKREYFNVALILPASKAKAPEYKPLFETKKNNA DYVTEBPENPNDPPTPEKONGNCEGCKKIFETPDPAAGKAVVLQGGGRVLEBMAH LRWKEFTNENNEQKFYLSNKEQOYICSAIRGNVLPDCCGSCYVRKPGADRVTGI FQKCEFTLPDOOSEKASIMPAQSIDIVEVECKRNNKREKPNQONKCLRSTREVI QDSDEFKRTTMTPTQPPAPFTSLQIQGVLCVLDVDSGWTGGRKLRLQDQKLEFL LQYEOGAMGVMAVFDASAAYKSELVQINSAEADALARS.LPTAASQGTSLSCGILRSAL FTVLKKRYPDGESEVLITLEDENMTISACEPEVQONGALHTVALIGSAAKELELS MTGLQRYASDQAENGLIDAFGALSSGNRAQSORSLQESQGLTLQNNEMNNTVV DSYQKDLPLTLEERKFLSPIPEGVSPGSGSDSLGKHKNKVAAPVPCATVAGMM KYSLOASSQITLTIVYSSRRSRTLPRVYTSKMKDCKPSPMVVYTKIHQGLPLIIG RAKVALIESENGKRTLELDNGAGADATKNDGILSYKFLYATANGKYSRWYALGG ASVTPRRAPPLMSQAMVLRWTEGELIKWNPFRINDKLDLQGVCKSRTASGSPF LNDLKEFNDSEQVWTTDLIPKEANSEEFVEKPRGIPPTNGTDLFLAQAADKTLKSK EISNAGVSLFLPPEAPETPEPPAPSLPCEPIOVNSTIGIHLKIMMKWLGETLOI SIA"

D	597	CCAGGAATGGGGCTACACTCGGATGGGGCGGTTTAATAGATGTCACAAACAGAGACAGAACTT	656
Q	549	CTACTTATCCAAATGGAGAATACAGCAGTAAAGATGTTTCAGCAGGTATTACTGTACAA	608
D	657	CTACTTATCCAAACAAAACAAACACAGTGAATATGTTGGCAGCATCCGGGACAA	716
Q	609	TGTTGTTAAAGAAAGTTCAGGAGGACGCTGTACACCAAAACATCCACTTCAATCAAAAGT	668
D	717	TGTACTACCAATGGCCMAAGGAGCGCTGTGCACCAAAACCGTCCGACGACGAGAGAGT	776
Q	669	AACAGGACTCTATGAAAAAAGATGTAGTTGTTTCCCAATCCGCCACAGAGAGAAAGC	728
D	777	AACGGGCTGTCCAGAGAGAGTGTAGTTTCATACCCTATCCCAAGCAGAGCAGAAAGC	836
Q	729	TTCATATATGTTTGCACACACATGTGATTCTATAGTTGAATTCGTAGACAAACAAACA	788
D	837	CTTATATATGTTTGCACAGAGCATGTATCTGTGGTTGAATTCGTAAAGAAAAAACCA	896
Q	789	CAACAAAGAAAGCTCCAAACAGCAAAATATAAAATATGCAATCTCCGACACATGGGAAT	848
D	897	CAATAAAGAAAGCTCCAAACAGCAAAACCAAAATATGCAATCTCCAGACATGGGAAT	956
Q	849	GATCCGATGTGAGAGACTTTAAAGAAAACCTCCTATGACAAACAGCAACCAAAATCC	908
D	957	GATCAAGATATCTGAAAGACTTTAAGAAAACCTCCTATGACAAACAGCAACCAACCGCC	1016
Q	909	CACCTTCTCATTTGCTGCGAGATTTGGACAAAGAAATTTGTGTGTTAGTCTTGACAAATCG	968
D	1017	CACCTTCTCATCTGCTGCAAGTGGACAGAAATCGTGTCTATGCTCTGCACAAAGTCTGG	1076
Q	969	AAGCATGGCAGCTGTACCCGCTCAATGAGTGAATCAAGCAGGCCACGTTTCTGCT	1028
D	1077	AAGCATACGTTGTGGGCGCTTCTTAAAGGAGCTGAATCAAGCAGGCCAACTCTTCTCTCT	1136
Q	1029	GCACAGATTTGAGCTGGGGGCTCGGGGTTGGAAATGGTGAATTTGGACAGTGGCCCATGT	1088
D	1137	GCACAGCTGTGAGCAGAGGGGCTGGGTTGGATGGTGGATTTGGACAGCGCTGCTTAATGT	1196
Q	1089	ACAAAGTGAACCTCATACAGATTAACACAGTGGCAGGTGACAGGAGACACTCTGCCAAAGAT	1148
D	1197	AAAAAGTGAACCTGCTCCAGATTAACACAGCGCCGGGAAAGACAGCGGCTTGGCCCGAGCT	1256
Q	1149	ACCTGCAGCAGCTTCAGAGGAGACGTTCATCTGCACGGGGCTTGCATCGCATTTTACTGT	1208
D	1257	ACCCACAGCAGCCTCGGAGAGAACGTTCATCTGCTGGGGGCTCCGCTCCGATTTACTGT	1316
Q	1209	GATTAGGAAGAAATATCCAACTGATGGATCTGAATTTGCTGCTGACGAGATGGGAGAA	1268
D	1317	GATTAAGAAAGAAATATCCCGACGACGAGGCTGTGAGATCGTTACGTGACGAGCGGGAGAA	1376
Q	1269	CAACACTATTAAGTGGGTGCTTTAACAGAGTCAACAAAGTGGTGCATATCCACACAGT	1328
D	1377	CAACAGCATTAAGCGCGCTGCTTCCGAGAGTGAAGCAAAACGACATATCCACACAGT	1436
Q	1329	CGCTTGGGGCCCTCTGCAGCTCAAGAACTAAGAGAGCTGTCCAAATATGACAGAGGTTT	1388
D	1437	CGCCCTTGGGGCCCTCGGACAGCAAGAGAGCTGAGAGAGCTGTGCAGATAGACAGGGGTTT	1496
Q	1389	ACACACATATGCTTCAGATTAACATTTAGAAACAAATGGCCCATATGATGCTTTTGGGGCCCT	1448
D	1497	GCAACACATATGCTTCAGATTAACATTTAGAAACAAACGGCTCATATGACGCTTTTGGGGCCCT	1556
Q	1449	TTCATCAGAAAAATGAGAGCTTCTCTAGAGGCTCCATCCAGCTGTGAAGTAAAGGATTAAC	1508
D	1557	TTCGTCGGGGAACAGAGCTCTCTCCAGGCTCCATCCAGCTGTGAAGTAAAGGATTAAC	1616
Q	1509	CTTCACAGAACGCCACTGATGATGAATGGCACACAGTATCTGACAGACACCGTGGGAAAGA	1568
D	1617	CTTCACAGAACCAAGAAATGATGAATGGCACAGGAGGTGGACACACACTGTGGGAAAGA	1676
Q	1569	CACCTTGTCTTATATCACTGAGCAACGAG----CTTCCCAATTCCTTCTCT--GGGA	1622
D	1677	CACCTTGTCTTATATCACTGAGCAACGAAAGTCTTCTTCCCAATTCCTTCTTCTTGGGGCT	1736

[illegible]

2700 CATTCCTGGCATTCATTTTAAAAATATGTCGAGTGCATGAGAACTGCAGCTGTC 2759
 2817 CATTCCTGGCATTCATTTTAAAAATATGTCGAGTGCATGAGAACTGCAGCTGTC 2876
 2760 AATGACCTGAGGCGATTTTGTGCATATTAATAATATCATTCCTTTTGTG 2819
 2877 GATGACCTGAGGCGATTTTGTGCATATTAATAATATCATTCCTTTTGTG 2930
 2820 AATGACCTGAGGCGATTTTGTGCATATTAATAATATCATTCCTTTTGTG 2879
 2931 AATGACCTGAGGCGATTTTGTGCATATTAATAATATCATTCCTTTTGTG 2990
 2880 TATGACTACA 2888
 2991 AATGCTAATA 2999

RESULT 6
AB017156
LOCUS AB017156 2937 bp mRNA ROD 10-NOV-1999
DEFINITION Mus musculus gob-5 mRNA, complete cds.
ACCESSION AB017156
VERSION AB017156.1 GI:3721911
KEYWORDS GOB-5.
SOURCE Mus musculus adult intestine goblet cell cDNA to mRNA.
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.
TITLE Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice
JOURNAL Biochem. Res. Commun. 255 (2), 347-351 (1999)
MEDLINE 2 (bases 1 to 2937)
REFERENCE Komiyama, T., Tanigawa, Y. and Hirohashi, S.
AUTHORS Direct Submission
TITLE Submitted (26-AUG-1998) to the DBJ/EMBL/GenBank databases. Tohru Komiyama, ERATO, JST, Genome Asymmetry Group, Doi Biosymmetry Project, Tokodai 5-9-9, Tsukuba, Ibaragi 300-2635, Japan
JOURNAL (E-mail: tkomedia@erato.tirc-net.co.jp, tel:81-298-48-1515, fax:81-298-47-8901)

FEATURES
source 1..2937
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_type="goblet cell"
 /dev_stage="adult"
 /tissue_type="intestine"
 /gene="gob-5"
 /gene="gob-5"
 /codon_start=1
 /protein_id="BAA33743.1"
 /db_xref="GI:3721912"
 /translation="MESTKSPVFLILHLEGVLSLSLIOLNNNGYEGIVAIADHVP
 EDALIOHKDWTOASPYLFAKFFKVALIIPESMAKPEYRPAKLETKNA
 DYLVSTTSPGIDDEPTHEIGAGEKIRILTPDLGKKTLOGPORFVEMAH
 FMYGVNEYNDDEKFLSKGRQAKNSALITGKNOVRCCOGSCITTKCVIDRVTG
 LYKDCNVEVPDQNEKASIMFNONINSVECTERKNNOEARPNONORCINRSWEY
 IOESBDFKQTPMTAPAPFSLIOGRIVCLVLDNSGSLNDRLRMQASRLF
 LLOTVFQSGSWKAVTFDSAAVYOSLEKINSAGRDLLKHLPTVSAGTSLCSGLRT
 AFYIKKKYTPDSEIYLLTDEDNITSCFDLVOSGAIHITVALGPAAMGSLVS
 KMTGCGQYSSPOYVNGIYDAFPAALSGNMAIAOHSIOLESRYVLONNOMNGSLV
 VDSVSGKDLFLITTTTHPTPIFIMPSVGNNGEITLDTTVAVILOYGCTAKGVFWK
 YSLQASSQTLTIVTSRAASATLPIITTVPVNKTGKPPSYVYASTROGASPLIR
 ASVTALIESNKGTVTLLELDNGAGADATKNGVSRFTAFDANKRTSVKIMALGCV
 TSDROAAPPKNRAATIDGMIEDGVRNMPPRETSYVODKOLCSRTSSGSEVATN
 VPAAPILPDLPPCCOITDLKASIOGONLVNLWTAVAGDDYDHGRASNTIIRSTSIYD
 LRHNTSLQVNTGLIPKASSEIEFEELGNTGNGTDIFIAIOAVDKSNLSSEI
 SNIRARVSPINQEPIDSPRCPDINSITPIPIHLKIMMKVLGSMOYTLGIH"

BASE COUNT 860 a 718 c 693 g 666 t

ORIGIN
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

Query Match 59.1%; Score 1764; DB 10; Length 2937;
 Best local similarity 77.5%; Pred. No. 0;
 Matches 2212; Conservative 0; Mismatches 615; Indels 29; Gaps 5;

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13


```
Db 975 GATGATCGTCTTAAACCGAATGAATACAGGACCGGCTTTCTGCTGCAGACTGTGGAG 1034
OY 1042 CTGGGGTCCCTGGGTTGGATGTGATGATTTGACAGTGTGCCATGTACAAAGTGAACTC 1101
Db 1035 CAGGGATCTCTGGGTGGGATGTGTGACCTTTGACAGTGTGCCATGTACAAAGTGAACTC 1094
OY 1102 ATACAGATTAACAGAGTGTGACAGTGTGACAGGACACACTGCGCAAAAAGATTACCTGACAGCT 1161
Db 1095 AAAACAGTTAAACAGTGTGCTGTACAGAGATCTGTATCAAGACACTTACCCACAGTATCT 1154
OY 1162 TCAGAGGAGGACGCTTCATCTGACAGCGGCTTTCATTCGGCATTTTACTGTGATTTAGGAAGAAA 1221
Db 1155 GAGAGGGAGGACATCTATGATCTCTGCGCTTGGACAGCATTTTACACTGATTAAGGAAGAG 1214
OY 1222 TATCCAACTGATGATCTGTAATTTGCTGTGCTGACAGATGGGGAAGACAACTATTAAGT 1281
Db 1215 TATCCAACTGATGATCTGTAATTTGCTGTGCTGACAGATGGGGAAGACAACTATTAAGT 1274
OY 1282 GGGTGTCTTAAACGAGTCAACAAAGTGTGCCATCATCCACACAGTGTGCTTTGGGGCC 1341
Db 1275 AACTGCTTTGACCTGTGAGGACAGAGCGGGCCATCATCATCATGAGTGTGCTTTGGGGCC 1334
OY 1342 TCTGACAGCTCAACAGTCAAGAGAGCTGTCCAAATATGACAGAGATTTACAGACATATGCT 1401
Db 1335 GCTGCGCTTAAAGAGCTTGAAGAGCTGTCCAAATATGACAGAGAGCTGTGACATATCT 1394
OY 1402 TCAGATCAAGTTCAGACAAATGAGCGCTTCATGATGCTTTGGGGCCCTTTTCATCAGGAAT 1461
Db 1395 TCGGATCAGAGTTCAGACAAATGAGCTGTCCAAATATGACAGAGAGCTGTGACATATCT 1454
OY 1462 GAGAGCTGTCTCAGCGCTTCATCAGCTTGAAGAGTAAAGGATTAACCTCCAGAAACAG 1521
Db 1455 GCGGCGCATGCTCAGAGCTTCATCAGCTTGAAGAGAGGAGGAGTAAATCTCCAGAAATAC 1514
OY 1522 CAGTGGATTAAGGACAGTGTGACAGAGAGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1581
Db 1515 CAGTGGATTAAGGACAGTGTGACAGAGAGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1574
OY 1582 ATACCTGTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1641
Db 1575 ATACCTGTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1634
OY 1642 GGTGCTTTGTAGTGTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1701
Db 1635 AATGCTTTTATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1694
OY 1702 AAGGTTGGACCTTGAATATACAGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1761
Db 1695 AAGGTTGGACCTTGAATATACAGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1754
OY 1762 ACGTCCCGTGTGCTCAATGCTACCTGCTCAATTAACAGTGTGCTCAATTAACAGTGTGCT 1821
Db 1755 ACGTCCCGTGTGCTCAATGCTACCTGCTCAATTAACAGTGTGCTCAATTAACAGTGTGCT 1814
OY 1822 GACACAGCAAAATTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1881
Db 1815 AACACAGGAGAAATTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1874
OY 1882 CCAATTCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1941
Db 1875 CCAATTCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1934
OY 1942 TTGGAACACTGATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2001
Db 1935 CTGGAATATCTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1994
OY 2002 AGGTATTTACAACTATATACAGAGATGTAGATACAGTGTAAAGTGTGGGCTGTGGGA 2061
Db 1995 AGGTATTTACAACTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2054
OY 2062 GGAGTTAAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2121
Db 2055 GGAGTTAACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2114
```

```
OY 2122 GCCTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2181
Db 2115 GCCTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2168
OY 2182 GATGTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2241
Db 2169 TATGTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2228
OY 2242 TCTGATGCTC---CAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298
Db 2239 ACCAATGCTCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2288
OY 2299 CTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2358
Db 2289 CTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2348
OY 2359 GATTTATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2418
Db 2349 GATTTATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2408
OY 2419 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2478
Db 2409 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2468
OY 2479 GCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2538
Db 2469 GCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2528
OY 2539 GATCTTTTCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2598
Db 2529 GATCTTTTCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2588
OY 2599 ATTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2658
Db 2589 ATTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2639
OY 2639 GAAAGCTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2718
Db 2640 GAAAGCTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2699
OY 2719 TTAATAATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2778
Db 2700 CTGAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2759
OY 2779 TTTGTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2838
Db 2760 TTTGTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2811
OY 2839 TGTATTTTACCTTCTGTTAGGGGGGAGATTAATA 2874
Db 2812 TGTATTTTACCTTCTGTTAGGGGGGAGATTAATA 2847
RESULT 7
AF127035 3204 bp mRNA PRI 11-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 2 (CACCC2)
DEFINITION mRNA, complete cds.
ACCESSION AF127035
VERSION AF127035.1 GI:5726288
KEYWORDS human.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3204)
AUTHORS Agnel,M., Vernalte,T. and Culouscou,J.M.
TITLE Identification of three novel members of the calcium-dependent
chloride channel (CaCC) family predominantly expressed in the
digestive tract and trachea
JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
```


D	b	1522	TAGTAATGCGCTGGATTGAGACAGCACTGTCTAATTTAGTAGTACAGTGGAAGAAAGCACGTT	1581
O	y	1575	GTTTTCTTATCACCCTGGACAACGACGCTCCGCCAAAATCCTTCCTGGGATVCCAGTGACA	1634
D	b	1582	CTTCTCATCACTAGAACAGTCTGCCTCCCGATTTTCTCTCTGGATVCCAGTGGAAC	1641
O	y	1635	GAACAGATGGGCTTGTGTACTGTGACAAAAMAAACCAAATAATGGGCTACCTCCAAATCCCAAG	1694
D	b	1642	AATAAATGAAAAATTTTCCACTGTGGATGACCACTTCCAAAATAGGCGCTATTCAGTAATTTCCAG	1701
O	y	1695	CATTGCTAAGGTTGGACCTTGGAATAATACAGTGTGC-----AAGAAGCTCATCAAACTT	1748
D	b	1702	AAC TGCAAAAGTGGGCACTTGGGCATACAAATCTTCAAGCCAANAAGGAACCGAABACATT	1761
O	y	1749	GACCCCTACTGTCAAGTCGCCGTGCCGTCCAATGCTACCCCTGCCCAATTAAGTAGCATTC	1808
D	b	1762	AAC TATTACAGTAACTCTTGAGACAGCAAAATTCCTGTGCTCCATACAGTAAGTAATGC	1821
O	y	1809	CAAAACACAAAGACACACAGCAAAATCCCCAGCCCTCTGGTAGTTATGCAAAATATTGC	1868
D	b	1822	TAAATTAATAAAGACGCTAAACAGTTTCCCAGCCCAATGAATGTTTATGCGAGAAATTC	1881
O	y	1869	CCAAGAGCCCTCCCAATTCCTCAGGGCCAGTGACACAGCCCTGATTAATCAGTAAGATGG	1928
D	b	1882	ACAAAGATATGTACTGTCTTCTTGAGCCCAATGAGTGTCTTCAATTAATACAGAAATGG	1941
O	y	1929	AAAAACAGTTTACCTTGGACATCTACTGATTAATGAGACAGTGTCTGATCTTAAGATGA	1988
D	b	1942	ACATACAGAAAGTTTGGAACTTTGGATATATAGTGTGAGGCGCTGATTTCTTCMAAGATGA	2001
O	y	1989	CGGTGTCTACTCAGGATTAATTCACAACTTAATGACAGAAATGGTAGATACAGTAAAGT	2048
D	b	2002	TGGAGTCTACTCCAGGATATTTTACACATATATACAGAAATGAGCAATTAACCTTAAGAT	2061
O	y	2049	GCGGGCTCTGGAGAGATTAATGACGACCCAGACAGAGAGATGAATCCCAAGACAGATGGAGC	2108
D	b	2062	TGGGGCTCATAGAGAGACAAACACTCCAGGCTTAAATTAACGGCTTCACATGTAATAGAC	2121
O	y	2109	ACTTACATATACCTGGCTGCAATTTAGAAATGATATAAATATACATATGATTCACCAAGACCTGA	2168
D	b	2122	CGGCTACATACAGGCTGGGTAGTGAACGGGAAATGAAAGCAAAACCCGCCAAGACTGA	2181
O	y	2169	AATTAATAAAGATGATGTTCAACACAAAGCAAGTGTGTTCACAGCAAGACATCCTCGGGAG	2228
D	b	2182	AATTTAT---GAGGATACCTGACAGCACTTTGGAGGATTTTCAACCGACAGCATCCGGAG	2238
O	y	2229	CTCATTTGTGGCTTCTGTATGTCCCAAAATGCTCCCATACTGATCTCTTCCCACTGGCCA	2288
D	b	2239	TGCATTTGTGTGTATCACAAATCTCCCAAGCCCTTCCCTGCTGACCAATACCCACAAAGTCA	2298
O	y	2289	AATACACGACCTGAAAGCGGAAATTAACGSGGGGCAATCTCATTAATCTGATCTGGACAGC	2348
D	b	2299	AATACACAGACCTTGATATGCCACAACTTATATAGG---ATAAGATTTATCTTCAATGGACAGC	2355
O	y	2349	TCTCTGGGATGATTAATGACCAATGAGGAACAGCTCACAAGATATATCATTCGTAATAGTACAG	2408
D	b	2356	ACCAAGAGATTAATTTTGTGTTGAGAAAGTCTCAAGSTTATATCATTAAGATTAAGTCAAG	2415
O	y	2409	TATTTCTGATCTCAGAGACAAAGTTCAATGATCTCTTCAAGTAATFACTACTGCTCAT	2468
D	b	2416	TATTTCTGATCTTAAGACACAGTTTATGATGTCTTTCACAGTAATTACTACTGATCTGTGC	2475
O	y	2469	CCCAAGAGAGCCAAACCTCAGAGAGCTTTTGGTTTAAACAGAAAAATTAATCTTTTGA	2528
D	b	2476	ACCAAAAGAGGCCCACTTCAAGAAAGCTTTGCATTTAAACAGAAAAATTTCTCAGAAGA	2535
O	y	2529	AAATGACACAGATCTTTTCAATGCTATTAAGCTGTGTAAGTGAAGTCTGTAATATCGA	2588
D	b	2536	AAATGCAACCCACATATTTTATGTCATTTAAAGTATAGATTAAGACAAATTTGACATCAA	2595
O	y	2589	AATATCCAACTTGACAGATATCTTTGTTATTTCTCCACAGACTCCGCAGAGACACC	2648

```

DB 2596 AGTATCCAAACATGTGCAAGAAAGTAACTTTGTTTATTCCTCCAGAAATTCGTATGACATTGA 2655
GY 2699 TAGTCTGATGAAACGCTGCTCTTGTCCCTPAATATTCATA 2699
DB 2696 TCTTACTCTCTACTCTACTCTACTCTGATGAAAAAGTCATA 2696

RESULT 18
AK000072 3221 bp mRNA PRI 22-FEB-2000
LOCUS AK000072
DEFINITION Homo sapiens cDNA FLJ20065 fis, clone COL01613, highly similar to
ACCION EICL BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.
VERSION AK000072.1 GI:7019922
KEYWORDS oligo,capping; fis (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_lib:COL clone:COL01613.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Carnalia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (stles)
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
TITLE Unpublished (2000)
JOURNAL 2 (bases 1 to 3221)
REFERENCE Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
AUTHORS Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo, Department
of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan, cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'- 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
Source
1..3221
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01613"
/clone_lib="COL"
/lssue_type="colon"
/note="Cloning vector pME18SFL3"
1..3221
/note="highly similar to EICL_BOVIN EPITHELIAL CHLORIDE
CHANNEL PROTEIN"

BASE COUNT 1105 a 600 c 634 g 882 t
ORIGIN 1
1
QueryMatch 43.7%; Score 1304; DB 9; Length 3221;
Best local Similarity 69.8%; Pred. No. 0;
Matches 1869; Conservative 0; Mismatches 785; Indels 24; Gaps
7;

GY 21 AGCATGAGGCCATTAAAGATTCGTGTTCAATTCATCTTCACCTCTAGAAAGGGGC 80
DB 39 AACATGGGGATTATTCAGAGGTTTGTTCCTCTTAGTGTGCGCTGCACACCAATC 98
GY 31 CCTGATTAATTCACATTCACGCTTAACAACAATGGCTATGAAGCATTTGTGTCAT 140
DB 99 ---AAATCTTCCTCATTAAGCTGAATATATGCTTTGAAGTATATGTCATTTAT 155
GY 111 CGACCCCATGGGCCAGGAAGATGAACACTCTCTCAACAAAATGAAGCATGTGACCA 200
DB 136 AGATCTTGTGTGTGCCAGGAAGATGAATAATATTAATCAAAATAGAGGATATGTACTAC 215
GY 201 GGCATCTCTGATATCTGTGTGAAGCTACAGAAAGCATTTTATTTCAAAAATGTGCCAT 260

```


Db 216 AGCTTACGTAACCTGTTGAGGACAGAAAAAGATTTTTCAAAAATGATCTAT 275
 Qy 261 TTTGATTCGGAAGATNGAAGCAAGAGGCTGACTATGTAGACCAAACTGTGACCTA 320
 Db 276 ATTAATTCCTGAGATGAGGAGGAAATCCTCACTACAAAAGGCGCAAAACCTGAAACCA 335
 Qy 321 CAAAAATGCTGATGTTGCTGCTGAGTCTACCTCCAGGTATGATGAACCTCTACAC 380
 Db 336 TAAACATGCTGATGTATAGTGTACACCACTACCTCCAGGTATGATGAACCTCTACAC 395
 Qy 381 TGAGCAGATGGGCAACTGTGTGAGAGAGAGGTGAAGAGATCCACCTCCAGTCTGATTCAT 440
 Db 396 CAAGCACTTCCAGAAATGTGAGAGAGAGAGGATATCATTCACCTCCAGTCTCT 455
 Qy 441 TGCAGAAAAAGTTACTCAATATGACACCAAGTATGAGGCACTTTTCCATGATGAGGCG 500
 Db 456 ACTTGAAGAAAAACAAATATATGAGACCAACGCAACTGTTTGTCCATGAGTGGCG 515
 Qy 501 TCATCTACGATGGGAGTATTTGACGAGTACATATATGATGAGAAATCTTACTTATCCAA 560
 Db 516 TCACCTCGGGTGGGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 575
 Qy 561 ---TGAAGATATACACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 617
 Db 576 GTCAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 635
 Qy 618 GAAGTGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 677
 Db 636 TAACTGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 695
 Qy 678 CTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737
 Db 696 GTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 755
 Qy 738 GTTTGCAACATGTTATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 797
 Db 756 GTTTATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 815
 Qy 798 AGCTCCAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 857
 Db 816 AGCTCCAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875
 Qy 858 TTCTGAGAGCTTAAAGAAACCACTCTATGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 917
 Db 876 TTCTGAGAGCTTAAAGAAACCACTCTATGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 935
 Qy 918 ATTGCTGACATTTGAGAAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 977
 Db 936 ATTGCTGACATTTGAGAAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 995
 Qy 978 GACTGTGAG 1037
 Db 996 GGGTGAAGAG 1055
 Qy 1038 TGAGCTGGGCTCTGGTGGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1097
 Db 1056 TGAAGATGAGATCTGGGTTGGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1115
 Qy 1098 ACTCATACATTAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1157
 Db 1116 GCTAATCCAAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1175
 Qy 1158 AGCTTCAAGAG 1217
 Db 1176 TCCCTGTGGAG 1235
 Qy 1218 GAAATAT---CAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
 Db 1236 GCTCATTCCTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1295
 Qy 1275 TATATAGGGGCTTTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1334

Db 1296 TGCAAGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1355
 Qy 1335 GGGGCGCTGTGAG 1394
 Db 1366 GGGAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1415
 Qy 1395 ATATGCTTCAGATCAAGTTGAGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1454
 Db 1416 TTATGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1475
 Qy 1455 AGGAATGAGAGCTGTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1514
 Db 1476 AGGAATGAGAGCTGTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1535
 Qy 1515 GAACAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1574
 Db 1536 TATGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1595
 Qy 1575 GTTTCTTATGACCTGTGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1634
 Db 1596 CTTTCTCATCAGATGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1655
 Qy 1635 GAACAGGAG 1694
 Db 1636 AATATGAG 1715
 Qy 1635 CATGCTGATGAG 1748
 Db 1716 AACTGTCAAGGAG 1775
 Qy 1749 GACCTGAGTGTGAG 1808
 Db 1776 AACTATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1835
 Qy 1809 CAAAGCAAGAG 1868
 Db 1836 TAAATATTAAGAG 1895
 Qy 1869 CAAAGGAG 1928
 Db 1896 ACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1955
 Qy 1929 AAAACAGTTACTTGAAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1988
 Db 1956 ACATACAGAAAGTTTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2015
 Qy 1989 CGGTGTACTCAAGGATTTTCAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2048
 Db 2016 TGAAGTCTACTCCAGTATTTTACAGCATATACAGAAATGAGATGATGATGATGATGATGATGATGAT 2075
 Qy 2049 GCGGCTCTGTGGAGGAGGATTTAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2108
 Db 2076 TCGGCTCATGAG 2135
 Qy 2109 ACTGTACATGAT 2168
 Db 2136 CCGCTTACTTACCGGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2195
 Qy 2159 AATTAATGAGGAT 2228
 Db 2196 AATGAT---GAGATCTCAGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2252
 Qy 2249 CTCATTTGTGCTTATGAT 2288
 Db 2253 TGCATTTGTGAT 2312
 Qy 2289 AATACCGAGCTGAG 2348
 Db 2313 AATACAGAGCTGAT 2369
 Qy 2349 TCCCTGGGAT 2408
 Db 2370 ACCAGGATATATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2429

	Do	2472	ACCAAGAGGAGCCAACTGCCAAGAAGAACCTTTGCATTTTAAACCACAAAATAATATTCAGAGA	2531
Cy	2539	AAAGGCGCAGCATCTTTTCATTGTCTATTCAGACGCTGGTCAATGAAGTCATGAATACAGA	2588	
Dg	2532	AAATGCAACCCACATATTATTATTTGGCATTAAAGATATGATATTAAGCATTTTGACATCAA	2591	
Gy	2589	AATATTCACAATTCGACAGACTATCTTTGTTATTCCTCCACACAGATCCGCCACAGA--C	2645	
Bd	2592	AGTATTCACAATTCGACAGAGTACTTGTATTCCTCCATACACCAATCTGATGACATTGA	2651	
Oy	2636	ACCTTAGTCTGATGAAACGCTGCTGCTTGTCTTCAATA	2683	
Dg	2632	TCTTACACACTACTCTACTCTACTCTACTCTACCTCCGATA	2689	
RESULT	10			
BTU36445				
LOCUS	BTU36445	2984 bp	mRNA	MAM 09-FEB-1996
DEFINITION	Bos taurus calcium-activated chloride channel mRNA, complete cds.			
ACCESSION	U36445			
VERSION	U36445.1	GI:1184065		
KEYWORDS				
SOURCE				
ORGANISM				
		cow.		
		Bos taurus		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;		
		Bovidae; Bovinae; Bos.		
REFERENCE		1 (bases 1 to 2984)		
AUTHORS		Cunhaingham,S.A., Awayda,M.S., Bublief,J.K., Ismailov,I.I.,		
		Arrate,M.P., Berdiay,B.K., Benos,D.J., and Fuller,C.M.		
		J. Biol. Chem. 270 (52), 31016-31026 (1995)		
TITLE		Cloning of an epithelial chloride channel from bovine trachea		
JOURNAL		66125078		
MEDLINE		2 (bases 1 to 2984)		
REFERENCE		Benos,D.J.		
AUTHORS		Direct Submission		
TITLE		Submitted (15-SEP-1995) Dale J. Benos, Physiology & Biophysics,		
JOURNAL		University of Alabama at Birmingham, BHSB 706, Birmingham, AL		
		35294, USA		
FEATURES				
source		Location/Qualifiers		
		1..2984		
		/organism="Bos taurus"		
		/db_xref="taxon:9913"		
		/cell_type="epithelial cells"		
		/tissue_type="trachea"		
		19..2730		
		/note="epithelial chloride channel"		
		/codon_start=1		
		/product="calcium-activated chloride channel"		
		/protein_id="AAC48511.1"		
		/db_xref="GI:1184066"		
		/translation="MVRPLVILFLTHLHPKMKSSWNLNNGDSIVIAINSYPE DEKIQIKKEVEASTLYLFHATRRRYEFNRVSLLIPMTKSKSEYLMPOQESDQAE VIVAPLKHGDDPYTLQVRCGRKGQYIHFTPEFLITNNLPYIGSGRAFEVEMAH, RWGFIDFNDGDOPYISRRNTIEATRSTHIGTNVNYCOGGSCITRPCRSQGL, YACATPTIPEKSQARSIIMFMOSLSHFTEFCETKINVAPIQNOMCNKGSWDVY MNSDPCSPMEPMNPOTPSLTSLKORVCYLVDKSGSMSSEDRPRMNDAELE FLIOITEKSGIAMPWDPAEVALRNNLTFTFDNVYENITANPDEANGSTCYRGAK AGQAQTIIQSOSTSGSEIILLTDDEDEHISCIEFYQSGVIHTTILGSAAKELET LSDMGGRHRFVANKDINGLATNFASIRSGSSTOOTLESRALATEKRWGVTP VDSITGDTFFVYVNIKTKEPILLQDRPKVKYKTSDEKDLNHSRLRPGLAETG TWYSLLNNHASPDQILVTTRARSPTPPPATAMSONTAYPSPVIVAVQSOQG FLPVLGINVAIITEDGHQVTELEMNGAGADLVKDGLSYRFTRYRNGRSKLKV HAERNNPARLSLNOPONKALYIPGYENKRTILNRPARYKDALAELIESRLFS GGSRTVGARPKNRPVLPNKTIIDLEAKREHDIOSWPARNYDLCKGANSTIINA SKSLDILOKPNDNLVNTSSLKFKEGESDENBEFKEPREIENTGNVITYAVOAIINEA NLNSEVSNIAQAIKIPMPDSDVPALGTJSAINALAFALAMILSTIV"		
BASE COUNT	1027 a	583 c	561 g	813 t
ORIGIN				
Query Match	31.3%	Score 933.4;	DB 4;	Length 2984;
Best Local Similarity	61.9%;	Pred. No. 1,8e-223;		

Matches 1635; Conservative 0; Mismatches 966; Indels 42; Gaps 8;

OY 13 AGATGTACAGCAATGGGCGCATTTAAGAGTCTGTGTTCATCTTGATCTTCACCTCTTA 72
 DB 7 ACATGCGCAAAATGGTCTGCTGACGTGCATCTGTCTTACTTGTGACCTCTGCG 66
 OY 73 GAAGGGGCCCTAGTAATTCATCTCAGCTGACGACAAACATAGGCGATATGAGGATTCG 132
 DB 67 CCTGG---AATGAAAAGTTCATGGTAAATTTGATTAACAAAGATATGATGGCATTTGTC 123
 OY 133 GTTGCAATCGACCCCAATGTGCGAGAAATGAAACACTCATTTCAACAAATTAAGAGCAT 192
 DB 124 ATTGCAATTAACCTTAGTGTGCGAGAGTGAATAAATCTATTCACAAATTAAGAGAAAG 183
 OY 193 GTGACCCAGCATCTCTGTATCTGTGTAAGCTACAGAAAGCATTTTATTTCAAAAT 252
 DB 184 GTAACGAAAGCTTCTACTACTCTGTTTCATGCGACCAAAAGAGATTATTTACAGAAAT 243
 OY 253 GTTCCATTTTATTCCTGCAAAACATGGAAGCAAAAGGCTGACTATGTGAGCCAAATCTT 312
 DB 244 GTAGCATTTTAAATTCATGAGCTGGAAGTCAAAATCTGACTTAAATGCGCAAAACAA 303
 OY 313 GAGACTCAAAATATGCTGATGCTGCTGAGTCTACTCTCCAGCTATATGATGA 372
 DB 304 GAATCATTTAGACAGGAGCAAGATCATAGTGTCTAATCTTACCTAAACATGAGATGAT 363
 OY 373 CCTCACTAGCAGATGGGCAACTGTGAGAGAAAGGTGAAGAGTCCACCTCACTCTCT 432
 DB 364 CCTTACACTTCATATGGAAGATGTGAGAAAAAGACATATATACATTTTACTCCA 423
 OY 433 GATTTCATTGCAAGAAAAATTTAGCTGAATATGACACAAAGTATGGCATTTGTCTAT 492
 DB 424 AACTCTCTGCTGACAAATTAATTTGCTATCTATGCTGCTCCGAGGCGAGACATTTGCTCAT 483
 OY 493 GAGTGGGCTCATCTGATGAGGAGATTTGAGAGTACAAATATGATGAGAAATCTCTAC 552
 DB 484 GAGTGGGCTCATCTGATGAGGAGATTTGAGAGTATGAGTATGAGGAGGAGGAGCATTTCTAT 543
 OY 553 TTATCCAAATGG---AAGATATACAGCAGTAAAGATGTTGACAGGATTAATCTGATCAAAAT 609
 DB 544 ATCTCAGAAAGAACATATGAGCAACAAAGATGTTCAACATATTTACTGTTACTAAT 603
 OY 610 GTAGTAAAGAGTGTGAGGAGGAGGAGCTGTACACCAAAAGTATGACATTTCAATTAAGTA 669
 DB 604 GGTATTTGCAAAATGCGAGGAGGAGCTATACAAAGGCGATGAGAGTACCTACACAG 663
 OY 670 ACAGAGCTCTATGAAAAAGAGTGTGATTTGTTCTCCATCCCGACGAGGAGAGAGCT 729
 DB 664 ACAGAGGCTGTATGAGGCAAAATGTACATTTCCAGAAAAATCCGAGCTGCAAGGAG 723
 OY 730 TCTATATGTTTGCACAAACATGTATGTTATGTTGAATTTCTGTACAGAACAAACACAC 789
 DB 724 TCCATCATGTTTATGCAAAAGTCTCATTTCTGTACGAAATTTTGTACAAAAAACACAC 783
 OY 790 AACAAAGAGCTCCAAACAGCAAAATCAAAATGCAATCTCCAGAGCAATGGAAGTG 849
 DB 784 AATGTGGAAGCTCCAAACCTCAAAACAAATGTGCAATGCGAAAGGACATAGGGATGTA 843
 OY 850 ATCCGTGATTTGAGAGATTTAAGAAACCACTCCATGACA-----ACAGAGCACA 903
 DB 844 ATCATGAACTCTACTGACTTTCAAAATACATCTCCATGACGAAATGATCCACGAGCT 903
 OY 904 AATCCACCTTCTCATTTGCTGATGAGTGGACAAAGAAATGTGTGTTTATGCTTGAACAA 963
 DB 904 CAATCATATTTTCAATTTGCTCAAGTCCAAACAGCAGTATGTTTGTGTTACTTGTATAA 963
 OY 964 TCTGGAACATGCGATGTGTAACCGCTCATGCTGATCAATCAAGAGGAGGAGGAGCTTTTC 1023
 DB 964 TCTGGAACATGCTTTCAGAAAGATGCTCTTCTTTCGAAATGAACAGCAGCAGCAATATATTC 1023
 OY 1024 CTGCTGACAGAGTGTGAGCTGGCTGAGGATGGATGGATGGATGGATGGATGGATGGATGG 1083
 DB 1024 TTGATTTCAATATTTGAAAAAGGATCTTGGTTGGATGGATTTGACAGTGTGCT 1083

OY 1084 CATGTACAAAGTGAACCTCATACAGATTAACAGTGGCAGTGAAGGAGACACTCGCCAAA 1143
 DB 1084 GAAATCCGAAATATCTACAAAAAATTAAGATGATTAATGTTATGAATAATATCACTGCA 1143
 OY 1144 AGATTAACCTGACAGAGCTTCAGAGGAGGAGCTCATATGCAAGGGGCTTGATGCGCATTT 1203
 DB 1144 AATGCTGCTCAAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203
 OY 1204 ACTGTATTTAGAGAGA---AATATCCACTGATGATGATGATGATGATGATGATGATGATGAT 1260
 DB 1204 CAGCATTTATTCAGAGTCAACAGAGCACTTCTGTTCTGAAATCTACTACTATACCGAT 1263
 OY 1264 GGGAGAGCAACACTAATAAGTGGTCTTTAAGAGGTCMAACAAAGTGGTGCATCAATC 1320
 DB 1264 GGGAGAGTAAATGAATAATACATCATGATGATGATGATGATGATGATGATGATGATGAT 1323
 OY 1321 CACACAGTGGCTTTGGGGGCTCTGAGCTCAAGAACTAGAGAGCTGTCCAAATGACA 1380
 DB 1324 CACACATTTGGCTGGGACCTTCTGCTGCAAGAACTGAGACATGTCAGATATGACA 1383
 OY 1381 GGAGTTTACAGACATATGCTTCAGATCAAGTTCAGAAACAAAGGCTCATGATGCTTTT 1440
 DB 1384 GAGGACATCGGTTTATGCTCAATTAAGACATA-----AATGCTTACTAATGCTTC 1437
 OY 1441 GGGGCTTTTATCAGGAATGAGCTGTCTCAGAGGCTTCATCCAGTTGAGATAG 1500
 DB 1448 AGTAGAATTTTATCTAGAAAGTGGCAGTATCTCAGACAGACTATTCAGTTGGAAAGCAA 1497
 OY 1501 GATTAACCTTCAGAACAGCAGCAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 1560
 DB 1498 GCTTGGGCAATTAAGAAAAAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1557
 OY 1551 GGAAGGACATTTGTTTATCACTGAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
 DB 1558 GGAATGACATTTCTTTTGTGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1617
 OY 1621 GATCCAGTGAAGACAGCA-----AGTGGCTTGTATGATGATGATGATGATGATGATGAT 1671
 DB 1618 GATCCAAAGAAAGAAATTAATAACCTCAGATTTCAAGAAAGATTAAGTAAATTTCTAT 1677
 OY 1672 ATGGCTTACCTCAAAATCCAGCAGTGTGAAGTTGGCAGCTTGGAAATACAGTCT--- 1727
 DB 1678 TCTGCTGCTTGAATACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1737
 OY 1728 -----GCAAGCAGCTCAAAACCTGACCTGATGATGATGATGATGATGATGATGATGAT 1782
 DB 1768 AATATCATGACAGTCTCAAAATCTATACAGTGAAGTACAGTCACTGAGCAAGAAAGTCT 1797
 OY 1783 ACCCTGCTCAATTAACAGTGAATTCACAAACAGAACAGACAGCAAGCAAAATTTCCAGAC 1842
 DB 1798 ACTACACCCCGATTAACAGTCAACAGCTCAGATGAGTCAAAATTAACAGACACATACCTGAC 1857
 OY 1843 CCTGCTGATGTTATGCAAAATTTGCGCAAGAGGCTCCCAATTTCTAGGCGCAGTGTCT 1902
 DB 1838 CAGAGTATGTTTATGCAAAAGTCAAGGCTTCTGCTGATGATGATGATGATGATGATGAT 1917
 OY 1903 ACAGCCGATTTGAATCAAGTGAATGAAAGAAAGATTAACCTTGAAGTCACTGATGATGAT 1962
 DB 1918 ACTGCAATTAAGAACTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1977
 OY 1953 GAGAGTGTGATGATTAAGATGAGTGTCTACTCAAGTATTTCAACATTAATGAC 2022
 DB 1978 GAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2037
 OY 2023 AGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2082
 DB 2038 GGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2097
 OY 2083 AGAGTATACCCAGCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2142
 DB 2098 AATTTAAGCAACACAGAACAAAGCCCTGTATATACAGGCTATATTTAAAAATGTTAA 2157

719 CGGAGAGGCTTCTAATATGTTGCAACAATGTTGATTCTATAGTTGAATTCCTGACAG 778
 760 CTGCAAGGAATCCATTATGTTATGCAAGCTCATCTGATGATGATTTTGTACAG 819
 779 AACAAACCAACAAGAGAGCTCCAAACAGCAAAATATAAATGCAATCTCCAGCA 838
 820 AAAAACAACACATACAGAGCTCCAAACCTACAAAACAAATGTGCAAGGCAAGCA 879
 839 CATGGGAATGATCGTGTATGAGGATTTAAGAAACCACTCTATAGCA-----A 892
 880 CATGGGATGATATCATGAATCTGTTGACTTTCAGATATCATCTCCATGACAGAAATGA 939
 893 CACAGCAACCAATCCACACTTCTCATTTGCTGAGATGAGCAAGAAATGTGTGTTAG 952
 940 ATCCACAGCATCTCATCTTATTTGATGCTCAAGTCCAAACAGCGGATGCTGTTGG 999
 953 TCCCTGACAAATCTGAGACATGGGACATGTAACCGCTCAATGACATGAATCAAGCA 1012
 1000 TACTTGATTAATCTGAAACATGCTGCAAGAACCGCTCTCTTCAAAATGCAAGCAG 1059
 1013 GCCAGCTTTCCTGCTGACAGAGTTGAGCTGGGGTCTGGGTTGGGATGGACATTGG 1072
 1060 CAGAACTATCTTGAATTCAGATTAATGAAAAGGATCTTTAGTTGGATGCTTACATTGG 1119
 1073 ACAGTGTGCTCCATGTACAAAGTGAATCATACAGATTAACAAGTGCAGTACAGGACA 1132
 1120 ACAGTGTGCTGAATTCAAATTCATCTACAAAGATTAACATGATTAATGTTTACCAA 1179
 1133 CACTGCGCCAAAAGATTACTGACAGACCTTCAGAGGAGACGTCATCTGACGCGCTTC 1192
 1180 AGATACCGCAAACTGCTCAAGTAGTAATGAGTGAATCTCAATTTTAAAGGCTCTCA 1239
 1193 GATGCGCATTTACTGATTTAGGAAGA---AATATCCATGATGATGATGAAATTTGTC 1249
 1240 AAGCAGAGATTCAGGCAATTAATCCACAGTGCAGACAGATCTTCTGTTGAATATATAC 1299
 1250 TCGTGAAGGAGGGAAGCAACACTATAAGTGGTCTTTAAACAAGTCAACAAAGTG 1309
 1300 TATTAATGATGGGAAGATTAATGAATTAATTCATGCTTGAAGATGTAACGAAGTG 1359
 1310 GGGCATCATTCACACAGTGGCTTTGGGCGCTCTGACACTCAAGACATAGAGAGCTGT 1369
 1360 GTGCAATTCATCCACACTTGTGAGACCTCTGCTCCAAAGAGAGAGACATTTGT 1419
 1370 CCAAAATGACAGAGGATTACAGACATATGCTTCAGATCAAGTTCAGAAATAGGCTCA 1429
 1420 CAAATATGACAGAGAGATTCGTTTTTTTCCCAATTAACATTA-----ACTGGCCTTA 1473
 1430 TTTGATGCTTTTGGGCGCTTTCATCAGAAATGAGAGCTGTCTCTGACGCTCCATCCAGC 1489
 1474 CTAATGCTTTTCAAGTAATTTATCTAGAAAGTGAAGCATCTACAGAGCTATTCAGT 1533
 1490 TTTGAGATGAGGATTAATCCCTCCAGAACAGCAGTGAATGATGACACAGTATGCTGG 1549
 1534 TGGAAAGCAAACTTGAATAATTAAGAAAGAAAGAGTAAGGCAAGGCAAGCTGTAG 1593
 1550 ACAGCAGCTGGGAAAGAGACATTTGTTTATCATCTACAGCAAGACGCTCCCAAA 1609
 1594 ACAGTACAGTTGGAATGACATTTCTTGTGTCAATGACAAATACAAAACAGAA 1653
 1610 TCCCTCTCTGGGATCCCACTGACAGAGCA-----AGTGGCTTTTGTAGTGACA 1660
 1654 TTTGCTCTCCAAATCCAAAGAAAGAAATTAATAAATCTGGAATTTCAAGAAAGATTAAGT 1713
 1661 AAAACACCAAAATGGCTACCTCCAAATCCAGGCAATTCCTAAGTTGGACACTTGAAT 1720
 1714 TAAATATTCATCTGCTGCTGCAAAATCTGTATTTCAAGACAGAGTACTTGGACTT 1773
 1721 ACAGTCT-----GCAAGCAAGCTCAAAACCTTGACCTGACGTCACGCTCGTG 1771
 1774 ACAGCTTCTAATATATATATGACAGCTCTCAATATGCTAAGTACAGTACAGTACAGT 1833

1772 CGTCAATGCTACCTCCCTCCCAATTTACAGTACTTCCAAAAGCAAGACACCAGCA 1831
 1834 CAGAAAGTCTCTACTATACCCCAAGTAAATGCAACACTCATCATGATCAACATACAGCA 1893
 1832 AATTCGAGCCCTCTGTAGTTATGCAATATTTGCCAAGAGGCTCCCAATTTCTCA 1891
 1894 ATTATCTGAGCCCAAGATGTTTATGACCAAGTCAAGTCAAGGTTTTTGGCTGTACTGG 1953
 1892 GGGGAGCTGTCACAGCCCTGATTTGAAATGATGAAAGAAACAGTTACTTGGAACTAC 1951
 1954 GAATCAGTATTAATGCAATTAATGAAACCAAGATGAGATCAATGAAATTTGGAGCTCT 2013
 1952 TGGATTAATGAGAGAGTGTCTGATGCTACTAAGATGAGCAGTGTCTACTCAAGTATTTCA 2011
 2014 GGGACATGAGTGTGAGTGTCTGATGCTGCAAGATGATGATCTCAAGTATTTCA 2073
 2012 CAATTTATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2071
 2074 CAGATTAATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2133
 2072 CAGCCAGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2131
 2114 CGGCTGAGCTTAATTTAATCAACACACAGACAAAGTCTTATATGTTCCAGGCTACGTTG 2193
 2132 AGAATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2191
 2194 AAAGGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2253
 2132 ACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2251
 2254 CTAATTAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2313
 2252 CAAATGCTCCCA---TACCTGATCTCTCCCACTGAGCAAGTCAACAGCTGAAGCGG 2308
 2314 CTCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2373
 2309 AAATTCACGGGGGAGTCTGATTAATCTGATGATGATGATGATGATGATGATGATGATG 2368
 2374 AGTTCAAGAGAG---ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2430
 2369 ATGGAACAGCTCACAAGTATATCTGATGATGATGATGATGATGATGATGATGATGATG 2428
 2431 AAGGAAGACCAACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2490
 2429 AGTCAATGATCTCTCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2488
 2491 ATTTCACATGAGCTTTAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2550
 2489 AGGAAGCTTTTGTAAACAGAAACATTAATTAATTAATTAATTAATTAATTAATTAAT 2548
 2551 AAGAAATTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2610
 2549 TTTGATGAGCTGTTGATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2608
 2611 TTTCACTCAAGCAGCAAGAGGCAATCTCATCTCAGAGGTTCTCATCAATTTGACAG 2670
 2699 TATCTTGTATTTCTCTCCAC 2629
 2611 CAATCAATTAATTCCTTAC 2691

RESULT 112
 AF115852
 LOCUS AF115852 2765 bp mRNA ROD 07-JAN-2001
 DEFINITION Mus musculus endothelial chloride channel (Clc2) mRNA, complete
 ACCESSION AF115852
 VERSION AF115852.1 GI:12043704
 KEYWORDS
 SOURCE house mouse,
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2765)
 AUTHORS Eble, R.C. and Pauli, B.
 TITLE A murine endothelial homolog of Lu-ECAM-1 associated with lung
 melanoma metastasis
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2765)
 AUTHORS Eble, R.C.
 TITLE Direct Submission
 JOURNAL Submitted (21-DEC-1998) Molecular Medicine, Cornell University, VMC
 C4-146, Ithaca, NY 14853, USA
 FEATURES
 source 1. 2765
 /organism="Mus musculus"
 /strain="C57/BL6"
 /db_xref="taxon:10090"
 1. 2765
 /gene="Clca2"
 17. 2725
 /note="putative metastatic addressin"
 /codon_start=1
 /product="endothelial chloride channel"
 /protein_id="A047626.1"
 /db_xref="GI:12043705"
 /translation="MVPGLQVLEFLTLHLNTESSNVHLNSNGEGLVAINPSVE
 DERLIPISKEMVTOASTLPEPASGRVYFNISILVPMWKSSEYLMPEKRESKAD
 VIVADPHLHGDPPYLOXGOCGRGOYIHFTPELLDNLRIYGRGRFVEMHML
 RMGFEDINDRPIYISKNTIEATRCASITGKRYHECORGCTRAARROSKPRL
 YEPKCTEIPDKIQTAGASIMFMNLSNVECTENNINAEPLNOKNCRSTWDVI
 KASAPONSPPMRGTETAPPEPTSLKSRWVCLYDKDSMKERLIRMNAAEL
 LYTOIVKESVGLVTEFDSAAHINLYLIKITSSDYOKITANLPQOATGGSICHGL
 AFOAITSDSOSTGSEIVLLDGEDNGISCFEASVRSGLIHTIALGSAARELET
 LSDWGGIRFPAKHHVSLIDAFSISSTSGSVQALLESKAFNVRAGMINSTVP
 VDSVGNTPFVITVYQKPELTIIDPKKRYITSDRDELNRSRLIOPGTAEGL
 TWTSITGTSQILITMTVTRASPTEPLIATLHMSQSTKQYSRHIVYRVSQGL
 PVLGNATVAILIEAHGQVLTLEMDNAGATVNDGITYTFDYGNGRTSLKRV
 QARKNARLSLRKNKSLYIPGYENGLVLPNPVQPEALIAVEDENRYSGS
 PTVSGAPDDGHAFPPSKVYDLEAEFIGDYIOLWTAPKVIDKGRANHYIIRVSG
 HPLGLOEDFNNAITLVNASSLIPKESGSETEKFEPEFKIANDTOLYIAIQYNEAGL
 TSEVSNIAQVKTSLSDISALCADISAIMTWGLAVINSLN"

BASE COUNT 832 a 656 c 635 g 642 t
 ORIGIN

Query Match 29.0%; Score 865.4; DB 10; Length 2765;
 Best Local Similarity 60.9%; Pred. No. 2,1e-206;
 Matches 1580; Conservative 0; Mismatches 976; Indels 39; Gaps 9;

440 AACTTGGATATCATATGAGACCCGAGCAGAGCTTGTCCATGAGTGGGCCCATCTCCGG 499
 511 TGGGAGATATTTACAGATACAAATATGATGAGAAATTTCTACTTATC---CAATGGAA 567
 500 TGGGAGATATTTATGATATTAAGTGAACCGGCTTTCTGATTTCTGAAAGAAACCT 559
 568 ATACAGACATAGATATGTTCCAGCAGGTATTCAGTACAAATGTACTAAAGAGTCCAG 627
 550 ATGAGCAACAAAGGTGTCTCCGACAGATCCAGGCAAGAAAGGTGTCCAGAGTGTAC 619
 628 GAGGACGCTGTACACCAAGATGACATTTCAATAAAGTAAACAGACTCTGTGAAAA 687
 620 AGAGGACGCTGTGACAAAGGCGTGGCGGCTGACTCCGAAGACAGCGCTGATGAAACC 679
 638 GAGTGTGATTTGTCTCCAAATCCCGGACAGAGAGAGAGAGCTCTTAATGTTGACAA 747
 690 AATGTACATTTATCCAGACAAATACAGACAGCTGGGGCTTCATTAATGTTATGCAA 739
 748 CATGTTATCTATAGTTGAATCTGTACAGAAACAAACCAACAAAGAGCTCCAAAC 807
 740 AACTCAATTTCTGTTGATTTTGCACAGAAATTAACCAATGACAGAAAGCCCAAC 799
 808 AAGCAATATTAATATCAATCTCCAGACACATGGGAAGTATCCGTGATTTGTGAGAC 867
 840 CTACAAACAAATGTCAATCGCAGAAACGCTGATTAATCAAGCGCTGTGCTGAC 859
 848 TTTAAGAAACCACTCTATG-----ACAACAGACACCAACCAATCCACTTCTGATTG 921
 860 TTTAGAAATTCCTCCATGAGAGAGAAACAAAGCGCTCTGCACTAATTTTCAATG 919
 922 CTGAGATTGGCAAGAAATTTGTGTTTGTAGTCTTCAAAATCTGAGACATGGCGACT 961
 920 CTCAAGTTCAGAAAGCGAGTGTGTGCTGTGATTAATCTGAAACATGAGAACAA 979
 932 GCTAACCGCCCATGCTACAGTAATCAAGACAGCGCAAGCTTTCCTGCTGACAGATTGAG 1041
 930 GAAAGCGCTTTATCGAATGAATCAAGACAGCAAGACTGTACTTAACTCAATTTGTGAA 1039
 1042 CTGGGCTCTGGTGGGATGTGACATTTGACAGTGTGCGCCATGTACAAAGTAACTC 1101
 1040 AAGAGTCTATGTTGGATTGTGACATTTGACAGCGCTGCCCATCAAAATTAATCTA 1099
 1102 ATACAGATTAACATGTCAGTGCAGAGACACACTCCGCCAAATATATCCGACAGCT 1161
 1100 ATAAATTAAGAGTATGATGACTACCAAAAGATCAACCGCAACCTCCCAACAGGCT 1159
 1162 TCAGAGGAGCTCCATCTGACGCGGCTGTGATCGGCAATTTACTGTGATTA---GGAAG 1218
 1140 ACTGTGGAACCTCAATTTGCGATGACACTCCAGCAGATTTTCAGCAATTAACCTCAGT 1219
 1219 AAATATCCACTGATGATTCGAATTTGTGCTGACAGATGGGAAAGCAACACTATA 1278
 1240 GACCAAGACCTTCGCTTGTGAGATCTGATGATGAGATGGGAAAGTAAAGGATA 1279
 1279 AGTGGCTCTTAAAGAGCTCAAAACAAAGTGTGCGATCATCCACAGCTGTTGGGG 1338
 1290 AGTCTCTCTTTAGGCGCTCTCTGACGCGGTGCTGATCATCAACCATGCTGTGGGG 1339
 1339 CCTCTGACGCTCAAGAACTAGAGAGCTGTCCAAATGACAGAGGTTTACAGACATAT 1388
 1340 CCTCTGCTGCCCGGACACTGTGACATCTGTACATGACAGAGAGGCTTCGTTTCTAT 1399
 1349 GCTTACATGATGATGACAAATGAGCTCTATGATGTTTGGGCGCTTCAATCAGGA 1458
 1430 GCCAACAAACATGT-----AAGAGCGCTTTCAGCGCTTTCAGTACGATTTCACTACA 1453
 1459 AATGAGCTGTCTCAGCGCTTCATCCAGCTTGAAGTAAGGATTAACCTCCAGAAC 1518
 1454 AGTGGCAGGTGTCCAGCAGGCTGTGAGATTTGAGAGCAAAAGCTTCAATGTACAGACA 1513
 1519 AGCCAGTGGATGAATGACACAGTGTATCTGACAGACACCGGTGGAAAGACACTTTGTT 1578


```

Db 1514 GGGGATGATTAACAGTACAGTACCTGTGACAGTACCTGGGACAGACAGTCTT 1573
Oy 1579 CTTATCACCCTGACAGACAGGCTCCCAATCTCTGAGGATCCAGTACAGAG 1638
Db 1574 GTTATCACCCTGACAGGATTAACAGGATTAATCATCTTCAAGATCCAAAGAAAA 1633
Oy 1639 CA-----AGTGGCTTTTGTAGTGCACAAAACACCAAAATGGCTTCTCAATC 1689
Db 1634 TATATTAACCTCAAGATTTCCAAATGATGTAATTAACATCCGGCTCTGACCTCAATA 1693
Oy 1690 CGAGGATGTTGAAGTGGACCTGGAAATACAGTCT---CCAAAGCAAGCTCACAAAC 1746
Db 1694 CCGGGACAGGACAGACAGAGTCTTGGATCTACAGATACAGGATACCAAGTCTCAGTTG 1753
Oy 1747 TTGACCTGACATGTCACAGCTCCGCTGCTCAATGTCACCTGCTCAATTAAGTACG 1806
Db 1754 ATTACAAATGACAGTACCCCTGACAGCAAGATCCACCATGGAACCATCTCGGTAC 1813
Oy 1807 TCCAAAGCAACAGACAGACACCAATTTCCAGCCCTCTGGTACTTTATGCAATAT 1866
Db 1814 GCTCAGATGAGTCAGAGCAGACCCAGTACCTAGCCGAGATGTTGTATGACGGTTC 1873
Oy 1867 GCGCAAGAGGCTCCCAATTTCTCAGGGCCAGTGCACAGCCCTGATTAATCAGTAA 1926
Db 1874 ACCCAAGGTTTTCCTGCTTCTAGAGGCAATGTCACAGATATGCAATACCACTGAAC 1933
Oy 1927 GGAAGAAACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1986
Db 1934 GACACATCAAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1993
Oy 1987 GAGGCTGCTCAAGTATTTACAACTTATGACAGATGATGATGATGATGATGATGAT 2046
Db 1994 GATGCACTCTACACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2053
Oy 2047 GGGGAGGCTGAGGAGGATTAAGCAGAGGAGGAGGATTAAGCAGAGGATGAT 2106
Db 2054 GGGGAGGCTGAGGAGGATTAAGCAGAGGAGGAGGATTAAGCAGAGGATGAT 2110
Oy 2107 GCACTGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2166
Db 2111 TCTTATATATATCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2170
Oy 2167 GAAATTAATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2226
Db 2171 GATGTCAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2230
Oy 2227 GGGTATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2283
Db 2231 GGGTATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2290
Oy 2284 GCGCAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2343
Db 2291 ACTTAAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2347
Oy 2344 AAGCTCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2403
Db 2348 AGGGGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2407
Oy 2404 ACAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2463
Db 2408 CAGCATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2467
Oy 2464 CTCATCTCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2523
Db 2468 CTGATCTCAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2527
Oy 2524 TTTGAAATGACAGATCTTTTCAATGATGATGATGATGATGATGATGATGATGAT 2583
Db 2528 ATAGCAATGACAGATCTTTTCAATGATGATGATGATGATGATGATGATGATGAT 2587
Oy 2584 TCAGAAATGACAGATCTTTTCAATGATGATGATGATGATGATGATGATGATGAT 2643
Db 2588 TCTGAGGCTCAAGATCTTTTCAATGATGATGATGATGATGATGATGATGATGAT 2647

```

```

Oy 2644 ACACCTAGTCTGAT 2658
Db 2648 GCACTGCTGCTGAT 2662
FEATURES
source
  1. 3137
  /organism="Mus musculus"
  /db_xref="locusID:80797"
  /db_xref="taxon:10090"
  /clone="MGC:5738 IMAGE:3583804"
  /tissue_type="mammary tumor. Metallothionein-TGF alpha
  model. 10 month old virgin mouse. Taken by biopsy."
  /lab_host="NCL CGAP_Mam1"
  /lab_vector="pCMV-SPORT6"
  /codon_start=1
  /product="chloride channel calcium activated 2"
  /protein_id="AA08147.1"
  /db_xref="GI:14198178"
  /translation="MPELOVLEFLLHLONTSSMYLNSNGEYVIAINPSVE
  DERLIPSTKEVNTQSTYLFASQGRVFRNISTLVPMTWRSKSEYLPKRESYKAD
  VIVADPHQDDPTLOYGCGDGRGQVIFHPNPLDNLRIYEPGRVYVHEAHL
  RMGVGDFRFAFNKIVSLIDAFSISSTSSVSQALSKAFNVAGAMINSTEY
  VDSVGNDFEYITWTYOKPELILDDPGKVIYSDPDDELINISARLIPYAEFG
  TWYYSINTEKSLITMTYETRAASPPMPLATAMSGSTAPYPSRMIVYVSGEFL
  PVLCANVAILIEABGHQVITLMDNGAGADTVKDGITTRFTFDYHNGRSLKLVY
  QARKNKLISLRKKSLYIPGYENGIVLNPDPDDEAIEATVDFNRYVSGS
  FTVSGAPDGDHARVFPSPKVTLEAEPIGDYIQLTWAPGKVLIDKGRHRYIIRVSG

```

```

BC008147 3137 bp mRNA ROD 12-JUL-2001
Mus musculus, chloride channel calcium activated 2, clone MGC:5738
IMAGE:3583804, mRNA, complete cds.
BC008147
BC008147.1 GI:14198177
MGC.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3137)
Strausberg, R.
Submitted (22-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalobedc.tmc.edu.
Villalob, D.K., Luna, R.A., Hale, S.M., Huiyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 10 Row: 9 Column: 5.
Location/Qualifiers
1. 3137
/organism="Mus musculus"
/db_xref="locusID:80797"
/db_xref="taxon:10090"
/clone="MGC:5738 IMAGE:3583804"
/tissue_type="mammary tumor. Metallothionein-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCL CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
20. 2728
/codon_start=1
/product="chloride channel calcium activated 2"
/protein_id="AA08147.1"
/db_xref="GI:14198178"
/translation="MPELOVLEFLLHLONTSSMYLNSNGEYVIAINPSVE
DERLIPSTKEVNTQSTYLFASQGRVFRNISTLVPMTWRSKSEYLPKRESYKAD
VIVADPHQDDPTLOYGCGDGRGQVIFHPNPLDNLRIYEPGRVYVHEAHL
RMGVGDFRFAFNKIVSLIDAFSISSTSSVSQALSKAFNVAGAMINSTEY
VDSVGNDFEYITWTYOKPELILDDPGKVIYSDPDDELINISARLIPYAEFG
TWYYSINTEKSLITMTYETRAASPPMPLATAMSGSTAPYPSRMIVYVSGEFL
PVLCANVAILIEABGHQVITLMDNGAGADTVKDGITTRFTFDYHNGRSLKLVY
QARKNKLISLRKKSLYIPGYENGIVLNPDPDDEAIEATVDFNRYVSGS
FTVSGAPDGDHARVFPSPKVTLEAEPIGDYIQLTWAPGKVLIDKGRHRYIIRVSG

```


Db 1401 GCCAACAAAGACCT-----AAACAGCCTTATGATGCTTTCAGTAGAATTTTCATCTACA 1454
 Oy 1459 AATGAGAGTGTCTCTCAGCGCTCCATCCAGCTTGAGAGTAAGGATTAACCTTCAGAAC 1518
 Db 1455 AGTGGAGAGGTCTCCAGAGAGGTCTGAGTGGAGAGCAAGAGCCTTCATGCTCAGAGCA 1514
 Oy 1519 AGCCAGTGAATGATGACAGTGAATGACAGCAGCCTGGAGAAAGACACTTTGTTT 1578
 Db 1515 GGGGAGTGAATTAAGGTACAGTACCTTGACAGTACCTGGACAGTCCGCAACAGAGTCTTT 1574
 Oy 1579 CTATATCACTGAGCAACGAGCCTCCCAATCTCTGGAGTCCAGTGGACAGAG 1638
 Db 1575 GTTATCACTGAGATGTAATAAAGAGAGATATTTTCAGATTCAGAAAGAAAAA 1634
 Oy 1639 CA-----AGTGGCTTTTGTAGTGACAAAACCAAAATGGCCTTACCTCAATC 1689
 Db 1635 TATACAACTCAGATTTCCAAAGATGATTAACATCCGCTGCTGAGACTTCAATA 1694
 Oy 1690 CAGAGCATTTCTAAGTGGTGGACTTGAATTAAGTCT---GCAAGCAAGCTCAGAAC 1746
 Db 1695 CCGGGGACTGACAGAGAGTACTTGGACTTACAGTACAGGGTACAGAGTCTCAGTTG 1754
 Oy 1747 TTGACCTGACTGCTCAGTCCGCTGCTCCATGCTACCTGCTCCATTTACAGTACT 1806
 Db 1755 ATTACATGACAGTACAGTCTCGAGCAAGATCCACATGAGAACACCTCGGCTACT 1814
 Oy 1807 TCAGAAAGCAAGACAGACCAAGCAATTTCCCAAGCCTCTGGTACTTTATGCAATAT 1866
 Db 1815 GCTCAGTGTAGTACAGACAGCCAGTACCTGAGCGGATGATGATGATGATGATGAT 1874
 Oy 1867 CGCAGAGAGCTCCCAATTTCTAGGGCCAGTGTACAGCAGCCTGATTTGATGATGAT 1926
 Db 1875 AGCCAGAGATTTTGGCTGCTGCTGAGGCAATGTACAGCCCTCATAGAGCTGAACAT 1934
 Oy 1927 GGAAGAAAGTGTACCTTGAATCTAGTATGATGATGATGATGATGATGATGATGAT 1986
 Db 1935 GGACATCAAGTACCTTGAATCTAGTATGATGATGATGATGATGATGATGATGAT 1994
 Oy 1987 GACGCTGTCTCAAGTATTTCAACATGATGATGATGATGATGATGATGATGATGAT 2046
 Db 1995 GATGCACTCTACAGAGTACTTTACAGATTTATCATGAAATGATGATGATGATGAT 2054
 Oy 2047 GTGGGGCTGTGGAGAGTAAACGACAGCAGAGGAGTATGATGATGATGATGATGAT 2106
 Db 2055 GTGGGTGTCCAGGCAAGAAACAAACAGACTGAGCTTAAGA---CAGAGAACAG 2111
 Oy 2107 GCACTGTACATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2166
 Db 2112 TCTTTATATATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2171
 Oy 2167 GAAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2226
 Db 2172 GATGTCCAGAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 2231
 Oy 2227 GCTTCATT---TGTGGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 2283
 Db 2232 GGGTGTGTACTGTGTGAGCGGCCCTGATGGCAGCAGGCTGTGTGTCCACCA 2291
 Oy 2284 GGGCAATACAGCAGTGAAGCGGAAATTCAGGGGGCAGTCTGATTAATCTGACTGG 2343
 Db 2292 AGTAAGTCCAGAGCTGAGAGCTGATGATGATGATGATGATGATGATGATGATGAT 2348
 Oy 2344 ACAGCTCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2403
 Db 2349 AGGGCCCTGAGAGGTTCTGACATGATGATGATGATGATGATGATGATGATGATGAT 2408
 Oy 2404 ACAAGTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2463
 Db 2409 CAGCATCTCTGTGATCTCCAGAGATTTTAACAATGCTACTTTAGTGAATGCTTCCAGT 2468
 Oy 2464 CTATCCCAAGAGCAACTGTGAGAGTCTTTTGTGTTAAACAGAAAAACATTTACT 2523
 Db 2469 CTGATACCTTAAGAGAGCTGTGCAAAAGAAATTTAAATTCAAACAGAAACTTTTAAA 2528

Oy 2524 TTGAAATGACAGAGATCTTTTCATTTGCTATTCAGGCTGTGATTAAGTCAATCTGAAA 2583
 Db 2529 ATAGCAATGAGCATTCAGCTCTTACATTTGATCCAGGCAAGACATGAAGCAGTCTCACC 2588
 Oy 2584 TCAGAAATATCCACATTTGACAGAGTATCTTTGTTTATTTCTCCACA 2630
 Db 2589 TGTGAGTCTCCACATTCGACAGGCTGTGCAAGCTTACTTCTTAGA 2635

Search completed: April 4, 2002, 03:19:06
 Job time: 61428 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:28:25 ; Search time 16681 Seconds
(without alignments)
1084.174 Million cell updates/sec

Title: US-09-049-696-19

Perfect score: 1683

Sequence: 1 AACAACTGCTGCATCATC.....AAATGCTAACCACTGGGTA 1683

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estlom:*
5: em_estlpl:*
6: em_estlba:*
7: em_estlro:*
8: em_estlov:*
9: em_estlc:*
10: qb_estl1:*
11: qb_estl2:*
12: qb_estl3:*
13: qb_estl4:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_tod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	939.4	55.8	2933	12 AK008659	AK008659 Mus muscu
2	928.4	55.2	2915	12 AK007466	AK007466 Mus muscu
3	692.8	41.2	712	10 AM009763	AM009763 wsr8b04.x
4	652.4	38.8	716	10 A1660234	A1660234 wsr8b02.x
5	648.4	38.5	734	10 A1660957	A1660957 wsr8b02.x
6	612.6	36.4	653	10 AM361534	AM361534 QV2-CT026
7	567.8	33.7	598	11 BE927204	BE927204 RCI-CT024
8	565	33.6	633	10 AM753451	AM753451 QV2-CT026
9	551.2	32.8	561	10 A1802756	A1802756 wsr8b04.x
10	516.8	30.7	528	11 BF083270	BF083270 RCI-CT024
11	492.6	29.3	625	10 AM361532	AM361532 QV2-CT026
12	489	29.1	504	10 A1802693	A1802693 wsr8b04.x

13	480	28.5	501	10 A1721275	A1721275 asr8b08.x
14	470.6	28.0	657	10 AM361521	AM361521 QV2-CT026
15	465.4	27.7	621	10 AM361520	AM361520 QV2-CT026
16	450	26.7	490	10 AM508854	AM508854 n12h10.9
17	449	26.7	652	10 AM753449	AM753449 QV2-CT026
18	445	26.4	450	10 AM009764	AM009764 wsr8b05.x
19	437.2	26.0	466	11 BG195179	BG195179 RST14358
20	434.4	25.8	561	10 AM361523	AM361523 QV2-CT026
21	434	25.8	455	11 BF083321	BF083321 RCI-CT024
22	432.4	25.7	496	10 AM753448	AM753448 QV2-CT026
23	421.8	25.1	935	10 BF578919	BF578919 QV2-CT026
24	411.8	24.5	906	11 BF581342	BF581342 602100726
25	400.4	23.8	654	10 AM361522	AM361522 QV2-CT026
26	379.4	22.5	911	11 BF578693	BF578693 602093175
27	369.4	21.9	974	11 BG952464	BG952464 602827322
28	368	21.9	392	10 A1687981	A1687981 wsr7f06.x
29	362.6	21.5	812	11 BG968878	BG968878 602836186
30	352.8	21.0	356	10 AM361528	AM361528 QV2-CT026
31	352.4	20.9	378	10 AA296955	AA296955 EST112726
32	336.8	20.0	646	10 AA691335	AA691335 vsl4h06.r
33	336	20.0	732	11 BG963719	BG963719 602831354
34	326.4	19.4	813	11 BG216984	BG216984 RST36685
35	326	19.4	479	10 AS41829	AS41829 vj01f01.r
36	323.8	19.2	469	10 BE136935	BE136935 ug57e12.y
37	322.2	19.1	335	10 A1721121	A1721121 asr3d08.x
38	303.2	18.0	455	10 AA871197	AA871197 vq32c11.r
39	298.2	17.7	450	10 AA734088	AA734088 vsl9c08.r
40	296.8	17.6	315	11 BF083396	BF083396 RCI-CT043
41	293.4	17.4	310	10 AA297150	AA297150 EST112734
42	291.2	17.3	528	11 BG965506	BG965506 602830872
43	289.8	17.2	629	10 AA518112	AA518112 v123c06.r
44	273.8	16.3	552	10 AA688953	AA688953 vsl04h07.r
45	263.4	15.7	507	10 AA623217	AA623217 v114d07.r

ALIGNMENTS

RESULT 1
AK008659
LOCUS
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210006G11, full insert sequence.
ACCESSION AK008659.1 GI:12842987
VERSION
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2933)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Methods in enzymology. 303, 19-44 (1999)
JOURNAL 99279253
MEDLINE 10349636
PUBMED
REFERENCE 2 (bases 1 to 2933)
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome research. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED
REFERENCE 3 (bases 1 to 2933)
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, K., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,


```
BASE COUNT      838 a      720 c      695 g      662 t
ORIGIN
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 day old"
```

Query Match	55.2%;	Score 928.4;	DB 12;	Length 2915;
Best Local Similarity	-76.7%;	Pred. NO. 4.1e-228;		
Matches 1210;	Conservative 0;	Mismatches 341;	Indels 27;	Gaps 5;

OY	1	AAACAAATGGTGGCATCATCATCAACAGTGGCTTTGGGGCCCTCTGACCTCAACAAACATAG	60
Db	1308	AGCAGACGGGGCCGCTCATCATCAACAGTGCCCTGGGACCGGCTGCCCTAAAGAGCTTG	1367
OY	61	AGAGACTGTCCAAATAATGACAGAGAG--TTTACAGACATATGCTTCAGATCAAGTTCCAGAC	119
Db	1368	AGCAGCTGTCCAAATAATGACAGAGAGCCCTGCAGACATTAATCTTCGGATCAGSTTCAGAC	1427
OY	120	AATGGCTCATTTGATGCTTTTGGGGCCCTTTTATCAGGAAATGAGAGCTGTCTTCAGGCC	179
Db	1428	AATGGCTTTTGTGATGTCTTTCGGAGACATCTCTCAGGAAATGCGGCATGGCTCAGAC	1487
OY	180	TCCATTCAGCTTGAGACTAAGGATTAACCTCCACGAACAGCAGCTGGATGATGATGGCACA	239
Db	1488	TCCTATTCAGCTGGAGAGACAGGGAGTTAATCTTCAGAAATTAACCATTGGATGAATGGCTCA	1547
OY	240	GTGATGCTGGACAGACCGGTGGGAAAGGACACTTGTGTTCTTATCAGCTGGACACAGCG	299
Db	1548	GTGATGCTGGACAGCTGGGTGGGCAAGACACTTGTGTTCTTATCAGCTGGACACAGCAT	1607
OY	300	CCCTCCCAATCTTCTCTGGGATCCCACTGGACAGACAAAGTGGCTTTGTAGTGAC	359
Db	1608	CCCTCTCAATATTATTACTGTGGATCCCAAGCGAGTGGAACAAATAGGTTTATATCTAAC	1667
OY	360	AAAAACACAAAATGGGCTGACCTCCAAATCCAGGACATGTAAAGTTGGACACTGGAAA	419
Db	1668	ACAAACCACTAAGGTGGCTACCTCCCAAGTCCAGGACGCGCTAAGSTTGGCTTTGGAAA	1727
OY	420	TACAGTGTGCAAGCAAGCTACCAAACTTGACCCCTGACTGTGACGTCCGTGCTCAAT	479
Db	1728	TACAGCATTCNAAGGAGACTCAGAGACTCTACCTTACTGTCACTCCCGTGCACACAAGT	1787
OY	480	GCTACCTGCGCTCCAATTTACAGTACTTCGAAACGACAAGACACACGACAAATTCCCC	539
Db	1788	GCTACACTGCTCTCTATTACAGGACCCCGGTAGTGAATMAACACAGGGAATTTCCC	1847
OY	540	AGCCCTCGGAGTTTATGCAATATTTCGCAAGGACCTCCCAATTCAGGGCCAGT	599
Db	1848	AGCCCTGTACAGCTGTATGCAAGCATTTCCGCAAGGAGCCTCCGCTATTTCTCAGGGCCAGC	1907
OY	600	GTCACAGCCCTGATTTGATCAGTGAATGGAACAAAGTTACCTTGGAACTACTGATAT	659
Db	1908	GTCACAGCCCTGATTTGATCAGTGAATGGAACAAAGTAAACCTTGGAAATTACTGATTAAC	1967
OY	660	GGAGCAGGTGCTGATGCTACTAAGGATGACGGTGTCTACTCAAGGATTTTCACAATTTAT	719
Db	1968	GGAGCAGGTGCGGATGTCACCAAGAAATGATGGTGTCTACTCAAGGTTTTTTACACCTTTT	2027
OY	720	GACACGAATGGTATATACAGTGTAAAGTCCGGGCTCTGSGAGAGATTAACGACCCAGGA	779
Db	2028	GATGCAAAATGATATACAGCTTAACCGCTTAAAAATATGGGCTCTGSGAGAGACTCACTTCACACGA	2087
OY	780	CGGAGAGTGAATCCCGACAGAGTGGAGCACTGTACATACCTGGCTGATTAAGAAATGAT	839
Db	2088	CAGAGAGAGCACTCCGAAAGAAACAGGCACTGTACATATGATGGCTGATTAAGAAATGAT	2147
OY	840	GAATATCAATGGAATCCACCAAGACTGAAATTAATTAAGGATGATGTTTCACACAGCA	899
Db	2148	GAAGTAAAGAAATGAACCCACCAAGCTCTGAAACATTA-----GTTATGTTTCACACAGCAAG	2201
OY	900	GTTGTTTTCACACAAACATCTCGGGGAGGCTATTTGTGGCTTCGATATGCC--CAAT	956
Db	2202	CTGTGCTTCACACAGACATCTTCAAGGGGAGCTTTGTGTGGCCACCAATGTCCCCGACACA	2261

QY	957	GCACCCATACATCTATCTCTCCACCTGGCCAAATCACCGACGACCTGAAAGCGGAATTCAC	1016
Dp	2262	GCCTCCATCTCTCACTCTTTCCACCCTCTCAAAATCAGTACGACGTGAAAGCCAGATCCAA	2321
QY	1017	GGGGGCGATCTCAATATCTGACTTGGACAGCTCTGGGGATGATTAATGACCAATGAAACA	1076
Dp	2322	GGGGCAGAACCTGGTGGATCTGACGAGGAGCGCTCCCTGGGGATATGACTACGACGGGAGA	2381
QY	1017	GCTCACAGATATCTGTAATGGAATAGTACAACTATCTTGATCTTCAGACCACTTGCAAT	1136
Dp	2382	GCTTCCACACTACATCTATCCGATGAGCCACCACTATCTTGATCTTCAGGACCACTTCAAC	2441
QY	1137	GAATCTCTTCACTGAAATACTACTGCTCTCATCCCAAGGAAAGCCAACTCGAGAGATC	1196
Dp	2442	ACCTCACTCCAACTGAACACTACCGGTCTTATCCCAAGAGGCGAGCTGAGGAATAC	2501
QY	1197	TTTTTGGTTTAAACCGAANAACATTACTTTTGAAATGGCACAGATCTTTTATTGCTATTT	1256
Dp	2502	TTTCAGATTGAATCGGAGGACCAACTTTTGGAAATGGCACGAAATCTTTATTGCTATTC	2561
QY	1257	CAGGCTTTGATTAAGGTCGATCGAATACAGAAATATCCACATTCGACAGATATCTTGG	1316
Dp	2562	CAGGCTTTGATTAAGTCCATTCGAAATAGAAATCTCCACATTCGACCGGTGTCTGTG	2621
QY	1317	TTTTTATCTCCACAGACTCGCGCCAGAGACACTAGTCTGATGAAAGCTGTGCTCTTGT	1376
Dp	2622	TTTCATCCCCGCTCAG-----GAGCGCCCAATTCGCCGAAGACTCAATCCCCCTGTGT	2672
QY	1377	CCCTAATTTCAATTCACAGACACCACTTCTGTGCATTCACATTTTAAATTAATGTGGAG	1436
Dp	2673	CCCTAATTCAGATCAACACACCACTTCTGTGCATTCACAGTCTGAATATATGTGGAG	2732
QY	1437	TGATATGAGAACTGAGAGTCTCAATAGCCTAGGGGTGATTTTTTGTGAGATAATAAAA	1496
Dp	2733	TGGCTAGGAGAAATGACGGTGAACAGTAGTTTGCACTGAATTTTCAGGCACGAANAATCAC	2792
QY	1497	TAAATCATTCATCTCTTTTGTGATTAATAATTTTCTTAAATATGATTTTATGACTTCTGT	1556
Dp	2793	CA-----GCTATTCCTTCACTCGAGAAATTTCTTAAATATGATTTTATGACTTCTGT	2844
QY	1557	TAGGGGCGATTAATACAAA 1574	
Dp	2845	TAGGGGCGGATTAATACAAA 2862	

RESULT 3
 AM009763/C
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

AM009763 712 bp mRNA EST 08-MAR-2000
 ws87b04.x1 NCI_CGAP.C03 Homo sapiens cDNA clone IMAGE:2504911 3'
 similar to TR:088826 088826 G0B-5 PROTEIN. ; mRNA sequence.
 AM009763
 AM009763.1 GI:5858541
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 712)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-rt@mail.nih.gov
 Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.,
 Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
www-bio.lnlnl.gov/dbtp/image/image.html

Db	535	CTGAGGAAGTCTTTTGGTTTAAACAGAACCACTACTTTTGAAAATGSCACAGATTCTT	476
OY	1246	TCATTGCATTCAGGCGGTGTATAAGCGCATTCGTAATTCAGAAATATCCAAACATGTGAC	1305
Db	475	TCATTGCATTCAGGCGGTGTATAAGCGCATTCGTAATTCAGAAATATCCAAACATGTGAC	416
OY	1306	GAGTATCTTGTGTTTATTCCTCCACAGACGCCGCGCAGAGACACCTAGTCCGTGATGAAACGT	1365
Db	415	GAGTATCTTGTGTTTATTCCTCCACAGACGCCGCGCAGAGACACCTAGTCCGTGATGAAACGT	356
OY	1366	CTGCTCCTCTTGTCCTAATATTCATATCAACAGCACACATTCCTGCGCATTTTAAAAA	1425
Db	355	CTGCTCCTCTTGTCCTAATATTCATATCAACAGCACACATTCCTGCGCATTTTAAAAA	296
OY	1426	TTATGTGGAAGTGTAGTAGAGACAGCTGCTGCTCAATAGCTAGGCGTGAATTTTGTCA	1485
Db	295	TTATGTGGAAGTGTAGTAGAGACAGCTGCTGCTCAATAGCTAGGCGTGAATTTTGTCA	236
OY	1486	GATTAATATTAATTAATCAATCAATCCCTTTTTTGGATTAATAAATTTCTAAANGTATTT	1545
Db	235	GATTAATATTAATTAATCAATCAATCCCTTTTTTGGATTAATAAATTTCTAAANGTATTT	176
OY	1546	TAGAATCTCTGTAGGGGGCGCATTAATCAATAGTATATGATATTAATCAATTAATGATTT	1605
Db	175	TAGAATCTCTGTAGGGGGCGCATTAATCAATAGTATATGATATTAATCAATTAATGATTT	116
OY	1606	CCTGTAGGGGGCGCATTAATCAATAGTATTTAGACTTCTGTAGGGGGCGATTAATAATAA	1665
Db	115	CCTGTAGGGGGCGCATTAATCAATAGTATTTAGACTTCTGTAGGGGGCGATTAATAATAA	56
OY	1666	ATGCTAACAACACTGGGCTA 1683	
Db	55	ATGCTAACAACACTGGGCTA 38	
RESULT	5		
LOCUS	A1660957/c		
DEFINITION	w120d08.x1 Soares_Dieckgraefe_colon_NHUC Homo sapiens cDNA clone IMAGE:2351151 3' similar to TR:088826 O88826 GDB-5 PROTEIN. ; , mRNA sequence.		
ACCESSION	A1660957		
VERSION	A1660957.1		
KEYWORDS	GI:4764540		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)		
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapbs.r@email.nih.gov This clone is available royalty-free through LLNL; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. Insert Length: 820 Std Error: 0.00 Seq primer: -400P from Gibco location/Qualifiers		
COMMENT	1..734		
FEATURES	source		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2351151"		
	/clone_id="Soares_Dieckgraefe_colon_NHUC"		
	/tissue_type="colon:mucosa from 5 ulcerative colitis patients"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: colon; Vector: p77j3d-pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'		

Query Match	Best Local Similarity	Score	DB	Length
Matches 707; Conservative	95.8%;	648.4;	DB 10;	734;
	0;	Pred. No. 4e-156;	Mismatches 25;	Indels 6; Gaps 4;
BASE COUNT	200 a	143 c	159 g	228 t
ORIGIN			4 others	
777	AGACGGAGAGGATPACCCAGCAGAGTGGAGCAGCTGTACATACCTGGCTGGATGGAGAT	836		
774	AGACGGAGAGGATPACCCAGCAGAGTGGAGCAGCTGTACATACCTGGATGGAGAT	675		
837	GATGAATATCAATGGAATCCACCAAGACCTGTAATTAATAGATGATGTCAACACAAAG	896		
674	GATG--AATACATGAATATCAACAGACATGAA--TAATTAAGATGATG-TCACACAAAG	619		
837	CAAGTGTCTTCAGCAGAAATCTCTGGAGGCTCATTTGTGGCTTGTATGTCCCAAT	956		
678	CAAGTGTCTTCAGCAGAAATCTCTGGAGGCTCATTTGTGGCTTGTATGTCCCAAT	561		
557	GCTCCATACCGATCTCTCCACCTGAGCCAAATACCGAGCCTTAAGGCGAAATTCAC	1016		
550	GCTCCATACCGATCTCTCCACCTGAGCCAAATACCGAGCCTTAAGGCGAAATTCAC	501		
1017	GGGGGACAGTCTTAATCTGACTTGGACAGCTCTGGAGGATGATTAATGACATGAACA	1076		
500	GGGGGACAGTCTTAATCTGACTTGGACAGCTCTGGAGGATGATTAATGACATGAACA	441		
1077	GCTCAAGTATATCATTCGATAGTAAAGTATCTTGATGATCAGACAAATTCAT	1136		
440	GCTCAAGTATATCATTCGATAGTAAAGTATCTTGATGATCAGACAAATTCAT	381		
1137	GAATCTCTCAAGTAAATCTACTGCTGCATCCCAAGAGACCACTGTGAGAGTTC	1196		
380	GAATCTCTCAAGTAAATCTACTGCTGCATCCCAAGAGACCACTGTGAGAGTTC	321		
1137	TTTTTGTTAAACCGAAGAAATCTTCTTTGAAAATGGCAGATCTTTTCTCTAT	1256		
330	TTTTTGTTAAACCGAAGAAATCTTCTTTGAAAATGGCAGATCTTTTCTCTAT	261		
1257	CAGGCTGTTGATAGTGTGATGTAATCAGAAATATCAATGTCAGAGATATCTTTG	1316		
260	CAGGCTGTTGATAGTGTGATGTAATCAGAAATATCAATGTCAGAGATATCTTTG	201		
1317	TTTATTCCTCCACAGACTCCGCGACAGACACTAGTCTGATGAAGACGTCTGCTCTGT	1376		
230	TTTATTCCTCCACAGACTCCGCGACAGACACTAGTCTGATGAAGACGTCTCTGT	141		
1337	CCTATATTCATATCAACAGACACACTCTGCTGATTCATTTTAAATTAATTTGGGAAG	1436		
140	CCTATATTCATATCAACAGACACACTCTGCTGATTCATTTTAAATTAATTTGGGAAG	81		
1437	TGGATAGGAGAACTCAGCTGTCAATAGCTTGGGCTGAAATTTTGTGCAGATAATATAA	1496		
90	TGGATAGGAGAACTCAGCTGTCAATAGCTTGGGCTGAAATTTTGTGCAGATAATATAA	21		
1497	TAAATCATTCATCTTTT 1514			
30	TAAATCATTCATCTTTT 3			

Email: asimpson@ludwig.org.br
 This sequence was derived from the RPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
<http://www.ludwig.org.br/scripts/gethtml2.pl?lcr=GV2&t2=QV2-CT0261261099-011-f01e43-1999-10-26&t4=1>
 Seq primer: puc 18 forward
 High quality sequence start: 37
 High quality sequence stop: 652.

1.653

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/decade="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESIES PCR (U.S. Letters Patent application No. 19
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      140 a      156 c      166 g      191 t
ORIGIN

```

Query Match	36.48;	Score 612.6;	DB 10;	Length 653;
Best Local Similarity	99.18;	Pred. No. 6.5e-147;		
Matches 637; Conservative	0;	Mismatches 4;	Indels 2;	Gaps 2

OY	217	ACAGCCAGTGGATAAATGAGCAAGATGATGTCGAGACACACCGTGGGAAAGACACTTTGT	27.6
Db	644	ACAGCCAGTGGTA -GAATGGCAGATGATCGTGACAGCACACCGTGGGAAAGACACTTTGT	58.6
OY	277	TTCTTATACCTGTGACACACGAGCCTCCCAATCTTTCTCTGGGATCCCAATGGACAGA	33.6
Db	585	TTTCTTGTACCTGTGACACACGAGCCTCCCAATCTTTCTCTGGGATCCCAATGGACAGA	52.6
OY	337	AGCAGAGTGCGTTTGATGATGACAAAAAACACAAATGGCGCTAGCTCAATCCAGAGCA	39.6
Db	525	AGCAAGGTGGCTTTGTATGATGACAAAAAACACAAATGGCGCTAGCTCAATCCAGAGCA	46.6
OY	397	TTTGCTAAGGTTGGCATTTGGAAATACAGTCTGCAGCAGCATCAAACTTGACCTGA	45.6
Db	465	TTTGCTAAGGTTGGCATTTGGAAATACAGTCTGCAGCAGCAGCATCAAACTTGACCTGA	40.6
OY	457	CTGTACAGTCCCGTGGCGCTCAATGCTACCGCTCTCCAAATTACAGTACTTCCAAAAAGA	51.6
Db	405	CTGTACAGTCCCGTGGCGCTCAATGCTACCGCTCTCCAAATTACAGTACTTCCAAAAAGA	34.6
OY	517	ACAAGGACACCAAGAAATTTCCCGCAGCCCTCTGGATGTTTATGCAAAATATTCGGCAAGAG	57.6
Db	345	ACAAGGACACCAAGAAATTTCCCGCAGCCCTCTGGATGTTTATGCAAAATATTCGGCAAGAG	28.6

RESULT:	7
FE927204/c	
LOCUS	BE927204
DEFINITION	598 bp mRNA
ACCESSION	U01602.1 Homo sapiens CDNA, mRNA sequence
VERSION:	BE927204
KEYWORD:	EST.
SOURCE:	human.
ORGANISM:	Homo sapiens

ORGANISM:

Eukaryota; Metazoa; Choriata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 598)

AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R., Nagai, M.A., da Silva, W. Jr., Zagó, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalhal, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brinhausen, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

KEYWORDS
20202663

COMMENT

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asampsoneludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/getthml2.pl?lt1=6t2-RC1-Cr0249-240>)
800-211-1106#3=2000-08-24&tt=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 598.

FEATURE SOURCE

•

/organism="Homo sapiens"

/db_xref="taxon:960

```
/clone_11b="CP0249"
/dev_stage="Adult"
```

/note="Organ: colon

Smal; A mini-library from ORESTES PCR (

716 - Ludwig Inst

into the puc18 vector. The mRNA and cDNA amplification

stringency conditions

BASE CONT	150 a	133 c	130 g
ORIGIN :			

100

Query Match	33.7%;	Score 567.8;	DB 11;	Length 598;
Best Local Similarity	99.3%;	Pred. No. 2.2e-135;		
Matches 591;	Conservative 0;	Mismatches 2;	Indels 2;	Gaps 2

OY	700	CAAGGATATTTACACACTTAATACACACCAAGGAGTGTATACAGTGTGTAAGAAAGGCGGGCTCTGG	759
Db	598	CAAGGATATTTACACACTTAATACACACGATGGTATGATACAGTGTGTAAGAAAGGCGGGCTCTGG	539
OY	760	GAGGAGTTAACGCACGACGAGAGAGTGTATACCCAGCAGAGTGTGAGCATCTGTACATAC	819
Db	538	GAGGAGTTAACGCACGACGAGAGTGTATACCCAGCAGAGTGTGAGCATCTGTACATAC	479
OY	820	CTGGCTGGATGAGAAATGATGTAATTAACAATGGAATCCACCAAGACCTGAAATTAATTAAG	879
Db	478	CTGGCTGGATGAGAAATGATGTAATTAACAATGGAATCCACCAAGACCTGAAATTAATTAAG	419
OY	880	ATGATGTTCAACACACAGCAAGCAAGTGTGTTTCAGCAGACATCTCTGGGAGGCTCATTTTGTGG	939
Db	418	ATGATGTTCAACACACAGCAAGCAAGTGTGTTTCAGCAGACATCTCTGGGAGGCTCATTTTGTGG	359
OY	940	CTTTCGATGTCGCAAAATCTCCATTAAGTCTCTTCCTCCACCTGGCGCAATATACGAGAC	999
Db	358	CTTTCGATGTCGCAAAATCTCTCCATTAAGTCTCTTCCTCCACCTGGCGCAATATACGAGAC	299

OY	1000	TGAGGCGGAAATTCACGGGGGAGCTCATTTAACTGCTGGAGACAGCTCCGGGGATG	1059
Db	298	TGAGGCGGAAATTCACGGGGGAGCTCATTTAACTGCTGGAGACAGCTCCGGGGATG	239
OY	1060	ATTATGACCATGGAGACAGCTCACAAAGTATATCATTTGGAATAAGTACAAAGTATCTTGATC	1119
Db	238	ATTATGACCATGGAGACAGCTCACAAAGTATATCATTTGGAATAAGTACAAAGTATCTTGATC	179
OY	1120	TCGAGACAAAGTCATGATCTCTTCAATGATGATATGCTGCTCATGCCAAGAAG	1179
Db	178	TCGAGACAAAGTCATGATCTCTTCAATGATGATATGCTGCTCATGCCAAGAAG	119
OY	1180	CCAACCTCTGAGGAGTCTTTTGTAAACGAGAAAACATTACTTTTGAANAATGGCACG	1239
Db	118	CCAACCTCTGAGGAGTCTTTTGTAAACGAGAAAACATTACTTTTGAANAATGGCACG	60
OY	1240	ATCTTTTCATTCGATTCAGGCTGTTAATGATCGATCGAATTCGAAATATATC	1294
Db	59	ATCTTTTCATTCGATTCAGGCTGTTAATGATCGATCGAATTCGAAATATATC	6

[illegible]

REFERENCES

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R. & de Souza, A.L. 2006. The evolution of the vertebrate eye. *Braz J Biol* 66(1): 1-17.

Huxley, C.D. 1989. *The evolution of the eye*. Oxford University Press, New York.

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

Phylum Chordata: Vertebrata: Euteleostomi: Eukaryota: Metazoa.

AUTHORS

Goldman, G.H., Carvalho, A.F., Matsumura, A., Baia, G.S., Simpson, D.H., Brunsteld, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with one
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
COMMENT
Contact: Simpson A.-J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL:
<http://www.ludwig.org.br/scripts/gethtml2.pl?ftl=QV2&2-QV2-Cf0261>
261059-011-005613-1999-10-26&4-1)
Seq. primer: puc 18 forward
High quality sequence start: 66
High quality sequence stop: 90

```

SOURCE
1. .033
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
136 a 151 c 161 g 185 t

```

Query, Match	Score	DB	Length
33.6%;	565;	10;	633;

Best Local Similarity	25.00	100.00	100.00
Matches	580	0	25
Conservative	0	0	25
Mismatches	25	0	0
Indels	0	0	0
Gaps	0	0	0

195 AGTAAGGATTAACTCCAGAACAGCCAGTGGATGATGCAATGGCAGTGTGACAGC 254

553

31A

[illegible][illegible]

0Y 315 CTCTGGGATCCCACTGGACAGAAAGCAAGCTGGCTTGGAGTGGACATTTATACACCAATTAAG 574

Db 432 CTCTGGATCCAGTGGACAGAAGCAAGCTGGCTTTGTAGTGGACAAAAACACCAAAATG 433

375 GCCTACCTCCAATCCAGGCATGCTAAGGTGGCACTTGGAAATACAGTCTGCAAGCA 434

432 GCCTACCTCCAATCCGAGCATGCTAAGTTGGCACTTGAATACAGTCTGCAAGCA 373

435 AGCTCACAACCTTGACCCCTGACTGTCACGTCGCCGTGCCATGCTACCCCTGCCCA 494

313

554

[illegible]

DB 032 A11ACAG1GAC11CCAAACCGAACCGACCGCCTT...AACCCTTCCTTTT

QY 555 TATGCAATATTCGCCAAGGAGCCICCCCAATTCACGGGCGAGIGICACAGCCCGAII 014

Db 252 TATGCAATATTTCGCCAAGAGCCTCCCAATTCTCAGGGCCAGTGTCAACAGCCCTGATT 193

615 GAATCAGTGAATGGAAAAACAGTTACCTTGGAACTACTGGATAATGGAGCAGGTCTGAT 674

Db 192 GAATCAGTGAATGGAACACAGTTACCTTGGAACTACTGGATAATGGAGCAGGTCTGAT 133

675 GCTACTAAGGATGACGGTGTCTCTACTCAAGGTATTTCACAACTTATGACACGAATGGTAGA 734

135 CCGACCTACCCCTCTACTTATGTATTCAGCAATTATGACCAAGATGGTGA 73

794

[illegible]

Lipid TACAGIGTAAAGAGCGGGCCTGGAAGAATTAACGCACCACCCGCCAACCAACCCCC

QY 7:35 CAGCA 799

381 AATACCTACTGCTCTCATCCCAAGGAAGCCCAACTCTGAGGAAGCTTTTGTGTTAAACCA 322

321 GAAACATTACTTTGAAATGGACACGATCTTTCTATGCTATTCAGGCTGTGGATAG 262
 122 GTGATCGAATCAGAAATATCCACATTCGACAGATATCTTGTATTCTCCACAG 133

07 1332 ACTCGGCGAGAGACACTAGTCCCTGGATGAAACGCTCTCTCTCTCTCTATATTTCATATC 139
 141 |||||
 142 |||||
 143 |||||
 144 |||||
 145 |||||
 146 |||||
 147 |||||
 148 |||||
 149 |||||
 150 |||||
 151 |||||
 152 |||||
 153 |||||
 154 |||||
 155 |||||
 156 |||||
 157 |||||
 158 |||||
 159 |||||
 160 |||||
 161 |||||
 162 |||||
 163 |||||
 164 |||||
 165 |||||
 166 |||||
 167 |||||
 168 |||||
 169 |||||
 170 |||||
 171 |||||
 172 |||||
 173 |||||
 174 |||||
 175 |||||
 176 |||||
 177 |||||
 178 |||||
 179 |||||
 180 |||||
 181 |||||
 182 |||||
 183 |||||
 184 |||||
 185 |||||
 186 |||||
 187 |||||
 188 |||||
 189 |||||
 190 |||||
 191 |||||
 192 |||||
 193 |||||
 194 |||||
 195 |||||
 196 |||||
 197 |||||
 198 |||||
 199 |||||
 200 |||||
 201 ACTCGCGAGAGACATCTAGTCCCTGGATGAAACGCTCTCTCTCTCTATATTTCATATC 142

Db 141 AAGAGCAACATTCCTGGATTCACATTTTAAAAATTTATGTGGGAAGTGGATAGGAGAACTG 82
 152 CAGCTGTCAAATAGCGCTTACGGCTGCAATTTTTCTCAGATAAATAATTAATCATTCATCTT 151

Sequence	Position
1502 TTTTTCATTATAAA	1528
1503 TTTTTCATTATAAA	1529
1504 TTTTTCATTATAAA	1530
1505 TTTTTCATTATAAA	1531
1506 TTTTTCATTATAAA	1532
1507 TTTTTCATTATAAA	1533
1508 TTTTTCATTATAAA	1534
1509 TTTTTCATTATAAA	1535
1510 TTTTTCATTATAAA	1536
1511 TTTTTCATTATAAA	1537
1512 TTTTTCATTATAAA	1538
1513 TTTTTCATTATAAA	1539
1514 TTTTTCATTATAAA	1540
1515 TTTTTCATTATAAA	1541
1516 TTTTTCATTATAAA	1542
1517 TTTTTCATTATAAA	1543
1518 TTTTTCATTATAAA	1544
1519 TTTTTCATTATAAA	1545
1520 TTTTTCATTATAAA	1546
1521 TTTTTCATTATAAA	1547
1522 TTTTTCATTATAAA	1548
1523 TTTTTCATTATAAA	1549
1524 TTTTTCATTATAAA	1550
1525 TTTTTCATTATAAA	1551
1526 TTTTTCATTATAAA	1552
1527 TTTTTCATTATAAA	1553
1528 TTTTTCATTATAAA	1554
1529 TTTTTCATTATAAA	1555
1530 TTTTTCATTATAAA	1556
1531 TTTTTCATTATAAA	1557
1532 TTTTTCATTATAAA	1558
1533 TTTTTCATTATAAA	1559
1534 TTTTTCATTATAAA	1560
1535 TTTTTCATTATAAA	1561
1536 TTTTTCATTATAAA	1562
1537 TTTTTCATTATAAA	1563
1538 TTTTTCATTATAAA	1564
1539 TTTTTCATTATAAA	1565
1540 TTTTTCATTATAAA	1566
1541 TTTTTCATTATAAA	1567
1542 TTTTTCATTATAAA	1568
1543 TTTTTCATTATAAA	1569
1544 TTTTTCATTATAAA	1570
1545 TTTTTCATTATAAA	1571
1546 TTTTTCATTATAAA	1572
1547 TTTTTCATTATAAA	1573
1548 TTTTTCATTATAAA	1574
1549 TTTTTCATTATAAA	1575
1550 TTTTTCATTATAAA	1576
1551 TTTTTCATTATAAA	1577
1552 TTTTTCATTATAAA	1578
1553 TTTTTCATTATAAA	1579
1554 TTTTTCATTATAAA	1580
1555 TTTTTCATTATAAA	1581
1556 TTTTTCATTATAAA	1582
1557 TTTTTCATTATAAA	1583
1558 TTTTTCATTATAAA	1584
1559 TTTTTCATTATAAA	1585
1560 TTTTTCATTATAAA	1586
1561 TTTTTCATTATAAA	1587
1562 TTTTTCATTATAAA	1588
1563 TTTTTCATTATAAA	1589
1564 TTTTTCATTATAAA	1590
1565 TTTTTCATTATAAA	1591
1566 TTTTTCATTATAAA	1592
1567 TTTTTCATTATAAA	1593
1568 TTTTTCATTATAAA	1594
1569 TTTTTCATTATAAA	1595
1570 TTTTTCATTATAAA	1596
1571 TTTTTCATTATAAA	1597
1572 TTTTTCATTATAAA	1598
1573 TTTTTCATTATAAA	1599
1574 TTTTTCATTATAAA	1600
1575 TTTTTCATTATAAA	1601
1576 TTTTTCATTATAAA	1602
1577 TTTTTCATTATAAA	1603
1578 TTTTTCATTATAAA	1604
1579 TTTTTCATTATAAA	1605
1580 TTTTTCATTATAAA	1606
1581 TTTTTCATTATAAA	1607
1582 TTTTTCATTATAAA	1608
1583 TTTTTCATTATAAA	1609
1584 TTTTTCATTATAAA	1610
1585 TTTTTCATTATAAA	1611
1586 TTTTTCATTATAAA	1612
1587 TTTTTCATTATAAA	1613
1588 TTTTTCATTATAAA	1614
1589 TTTTTCATTATAAA	1615
1590 TTTTTCATTATAAA	1616
1591 TTTTTCATTATAAA	1617
1592 TTTTTCATTATAAA	1618
1593 TTTTTCATTATAAA	1619
1594 TTTTTCATTATAAA	1620
1595 TTTTTCATTATAAA	1621
1596 TTTTTCATTATAAA	1622
1597 TTTTTCATTATAAA	1623
1598 TTTTTCATTATAAA	1624
1599 TTTTTCATTATAAA	1625
1600 TTTTTCATTATAAA	1626
1601 TTTTTCATTATAAA	1627
1602 TTTTTCATTATAAA	1628
1603 TTTTTCATTATAAA	1629
1604 TTTTTCATTATAAA	1630
1605 TTTTTCATTATAAA	1631
1606 TTTTTCATTATAAA	1632
1607 TTTTTCATTATAAA	1633
1608 TTTTTCATTATAAA	1634
1609 TTTTTCATTATAAA	1635
1610 TTTTTCATTATAAA	1636
1611 TTTTTCATTATAAA	1637
1612 TTTTTCATTATAAA	1638
1613 TTTTTCATTATAAA	1639
1614 TTTTTCATTATAAA	1640
1615 TTTTTCATTATAAA	1641
1616 TTTTTCATTATAAA	1642
1617 TTTTTCATTATAAA	1643
1618 TTTTTCATTATAAA	1644
1619 TTTTTCATTATAAA	1645
1620 TTTTTCATTATAAA	1646
1621 TTTTTC	

BF083270	528 bp	mRNA	EST	18-OCT-2000
LOCUS				
DEFINITION	RC1-CT0249-110900-214-b01 CT0249	Homo sapiens	cDNA, mRNA sequence	
ACCESSION	BF083270			
VERSION	BF083270.1	GT-10877100		

1 SOURCE: 1
2 ORGANISM: human.
3 Homo sapiens
4 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
5 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
6
7 REFERENCE 1 (phases 1 to 528)

Morgan, M.R., Carvalho, J., Zagury, A.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsushima, A., Bala, G.S., Simpson, Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hara, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
COMMENT?
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics

Brazil
Rua 1001, Anicão, Ladeira 107, 4 andar, 01309-020, São Paulo
Tel.: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

(<http://www.ludwig.org.br/scripts/gethtml2.pl?l=6&2=RC1-CY024>)
900-214-b01&t3=2000-09-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 33

```
source
1.528
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="CT0249"
```

Small; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No.

716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT	158 a	113 c	127 g	130 t
ORIGIN				

Query Match:	30.7%	Score 516.8:	DB 11:	Length 528:
Best Local Similarity:	98.7%	Pred. No. 2, 9e-122:		
Matches 521:	Conservative	0:	Mismatches 7:	Indels 0:
				Gaps 0:

QY	640	CCGGAACTCTGGAATTAAGGACAGCAGTGTGTTGCTTAAGGAATGACGGTGTACT	699
Db	1	CGTGAATATTTCTGTGATTAATGAGCAGAGTGTCTGATGCTACTAAGGATGACGGTGTACT	60
QY	700	CAAGGATTTTCAACAATTATGACACGAANTGGTAGATACAGTGTAAAGTGCGGGCTCTGG	759
Db	61	CAAGGATTTTCAACAATTATGACACGAANTGGTAGATACAGTGTAAAGTGCGGGCTCTGG	120
QY	760	GAGGAGTTAAGCGCCAGACAGCGAGATGATATCCCGACGAGTGGAGCTGTACTATAC	819
Db	121	GAGGAGTTAAGCGCCAGACAGCGAGATGATATCCCGACGAGTGGAGCTGTACTATAC	180
QY	820	CTGGCTGTGATTTGAGAAATGATGAAATTAACAAATGAAATCCACCAAGACCTGAAATTAATAG	879
Db	181	CTGGCTGTGATTTGAGAAATGATGAAATTAACAAATGAAATCCACCAAGACCTGAAATTAATAG	240
QY	880	ATGATGTTCAACACAAACCAAGTGTGTTTCACAGAAACATCTCGGGAGGCTCATTTGTGG	939
Db	241	ATGATGTTCAACACAAACCAAGTGTGTTTCACAGAAACATCTCGGGAGGCTCATTTGTGG	300
QY	940	CTTCTGATGTCCCAATGCTCCCTACCTGATCTCTTCCCACTGGGCAANTACCGGACC	999
Db	301	CTTCTGATGTCCCAATGCTCCCTACCTGATCTCTTCCCACTGGGCAANTACCGGACC	360
QY	1000	TGAAGCGGAAATTCAGCGGGGACGTCTCATTATCTGACTTTGGACAGCTCTCGGGGATG	105
Db	361	TGAAGCGGAAATTCAGCGGGGACGTCTCATTATCTGACTTTGGACAGCTCTCGGGGATG	420
QY	1060	ATTATGACCATGGAACAGCTCCACAAGATATTCATTCCAAATAGATACAAAGTATCTTGATC	111
Db	421	ATTATGACCATGGAACAGCTCCACAAGATATTCATTCCAAATAGATACAAAGTATCTTGATC	480
QY	1120	TCAGAGCAACTGCAATGAATCCTTCAAGGAATTAATAGTCTCTCA	1167
Db	481	TCAGAGCAACTGCAATGAATCCTTCAAGGAATTAATAGTCTCTCTCA	528

RESULT	11
LOCUS	AM361532
DEFINITION	AM361532 625 bp mRNA EST 04-FEB-2000
ACCESSION	U02-CT0261-261099-011-e03 CT0261 Homo sapiens cDNA, mRNA sequence
VERSION	AM361532
KEYWORDS	AM361532.1 GI:6866286
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryotic Metazoa: Chordata: Craniala: Vertebrata: Euteleostomi;
TITLE	Mammalia: Eutheria: Primates: Catarrhini: Homiidae; Homo.
JOURNAL	1 (bases 1 to 625)
COMMENT	HCCP http://www.ludwig.org.br/ORESTES . The FAPESP/LICR Human Cancer Genome Project Unpublished (1999) Contact: Simpson A.J.G.

project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?PL1-QV2&t2-QV2-CT0261>)

261099-011-e03&t3-1999-10-26&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 93

High quality sequence stop: 608.

Location/Qualifiers

Query Match	29.3%	Score 492.6	DB 10	Length 625
Best local Similarity	99.2%	Pred. No. 4.9e-116		
Matches 495	Conservative	0	Mismatches 4	Indels 0
				Gaps 0

Query Match	29.3%	Score 492.6	DB 10	Length 625
Best local Similarity	99.2%	Pred. No. 4.9e-116		
Matches 495	Conservative	0	Mismatches 4	Indels 0
				Gaps 0

QY	2.1	AAAAAGGGGGCCATCATATCCACACAGTCGGTTTGGGGCCCTCTGCAGCTCAGAACTAG	60
Db	127	AACATAGTGGTGGCCCTCATGACACAGTCGGTTTGGGGCCCTCTGCAGCTCAGAACTAG	186
QY	5.1	AGGAGCTGTCCAAAATGACAGAGAGGTTTACAGACATATGCTTCAGATCAAGTCCAGACA	120
Db	137	AGGAGCTGTCCAAAATGACAGAGAGGTTTACAGACATATGCTTCAGATCAAGTCCAGACA	246
QY	12.1	ATGGCCCTATTTGATGCTTTTGGGGCCCTTTCATCATGAGAAATGGAGCTGTCTCCAGGCT	180
Pb	247	ATGGCCCTCATTTGATGCTTTTGGGGCCCTTTCATCATGAGAAATGGAGCTGTCTCCAGGCT	306
QY	19.1	CCATCCAGCTTTGAGAGTTAAGGATTTAAACCTCCAGAACAGCCAGTGGATGAATGGCCAG	240
Pb	337	CCATCCAGCTTTGAGAGTTAAGGATTTAAACCTCCAGAACAGCCAGTGGATGAATGGCCAG	366
QY	24.1	TGATGCTGGAGACAGACCGTGGGAAAGACACTTTGTTCTTATACCTGGACAAAGCAGC	300
Pb	357	TGATGCTGGAGACAGACCGTGGGAAAGACACTTTGTTCTTATACCTGGACAAAGCAGC	426
QY	30.1	CTCCCCAAATCTCTCTGGGATCCAGTGGACAGAGCAAGAGTGGCTTTGTATGTATGACA	360
Pb	437	CTCCCCAAATCTCTCTGGGATCCAGTGGACAGAGCAAGAGTGGCTTTGTATGTATGACA	486
QY	36.1	AAAACACCCAAAATGGCCCTACCTCCAAATCCAGGACTTGTCTAAGGGTGGGCACTTGGAAAT	420
Db	437	AAAACACCCAAAATGGCCCTACCTCCAAATCCAGGACTTGTCTAAGGGTGGGCACTTGGAAAT	516
QY	42.1	ACAGCTCTGCAAGCAAGCTCACAAACCTTGAACCTGTACGTACCTCCCGTGCCTCAATG	480
Pb	517	ACAGCTCTGCAAGCAAGCTCACAAACCTTGAACCTGTACGTACCTCCCGTGCCTCAATG	606
QY	48.1	CTACCCCTGCCCTCCAAATTAC 499	
Pb	637	CTACCCCTGCCCTCCAAATTAC 625	

RESULT	112
AIR02693/c	
LOCUS	*
DEFINITION	wf77c01.x1 Soares_Dieckgreife.colon.NHUC Homo sapiens CDNA clone IMAGE:2350848 3' similar to IR:088826 OBB826 GDB-5 PROTEIN. ; , mRNA sequence.
ACCESSION	AIR02693
VERSION	AIR02693
KEYWORDS	AIR02693.1 GI:5368165 EST 18-DEC-1999

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 504)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgap@remail.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (infoimage.llnl.gov) for further information.
Insert length: 727 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 392.
FEATURES
Source
1..504
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2350848"
/clone_1lb="Soares, Dieckgraefe, colon NHUC"
/tissue_type="colonic mucosa from 5 ulcerative colitis patients"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site: 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCATCTGAGTGGAGCGGCGCTACTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieck@um.wustl.edu); colonic mucosa represents a range of disease involvement from mild cryptitis to severe ulceration, fibrosis, and degeneration. Library constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 150 a 85 c 103 g 163 t 3 others
ORIGIN
Query Match 29.1%; Score 489; DB 10; Length 504;
Best Local Similarity 99.4%; Pred. No. 4.1e-115;
Matches 489; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1023 AGCTCATTAATCTGACTGGAGAGCTCTGGGAGATTAATACATGAGACGCTCAC 1082
|||||
Db 504 AGCTCATTAATCTGACTGGAGAGCTCTGGGAGATTAATACATGAGACGCTCAC 445
QY 1083 AAGTATATCATTCGATTAAGTACAAATATCTGATCTCGAGACAAATGTCATTAATCT 1142
|||||
Db 444 AAGTATATCATTCGATTAAGTACAAATATCTGATCTCGAGACAAATGTCATTAATCT 385
QY 1143 CTTCAGATTAATCTGACTCTGCTCATCCCAAGGAAAGCACTCTGAGAGAGCTCTTTTG 1202
|||||
Db 384 CTTCAGATTAATCTGACTCTGCTCATCCCAAGGAAAGCACTCTGAGAGAGCTCTTTTG 325
QY 1203 TTTAAACCGAAACATTAATCTTTGAAATGGACAGATCTTTTCAATGCTATAGGCT 1262
|||||
Db 324 TTTAAACCGAAACATTAATCTTTGAAATGGACAGATCTTTTCAATGCTATAGGCT 265
QY 1263 GTTGATTAAGTCTGATTAAGTACAAATATCCACATTCAGAGATTAATCTTTGTTAT 1322
|||||
Db 264 GTTGATTAAGTCTGATTAAGTACAAATATCCACATTCAGAGATTAATCTTTGTTAT 205
QY 1323 CCTCCACAGACTCCGCGAGAGACCTAGTCTGATGAAGAGCTCTGCTCTTCTTAAT 1382
|||||
Db 204 CCTCCACAGACTCCGCGAGAGACCTAGTCTGATGAAGAGCTCTGCTCTTCTTAAT 145
QY 1383 ATTCAATTAAGACAGCACTCTCTGCTGATTCACATTTTAAATTAATGAGAGTGGATA 1442
|||||
Db 144 ATTCAATTAAGACAGCACTCTCTGCTGATTCACATTTTAAATTAATGAGAGTGGATA 85

QY 1443 GGAGAACTGCAGCTGTCATAGCTAGGCTGAATTTTGTGACATAATAATAATATC 1502
|||||
Db 34 GGAGAACTGCAGCTGTCATAGCTAGGCTGAATTTTGTGACATAATAATAATATC 25
QY 1503 ATTCAATCTTTT 1514
|||||
Db 34 ATTCAATCTTTT 13
RESULT 113
A1721275 501 bp mRNA EST 10-JUN-1999
A1721275/c
LOCUS 2
DEFINITION as82h08.x1 Barstead colon HPLR87 Homo sapiens cDNA clone
IMAGE:235263.3, similar to SW-620.BOVIN P54281 EITHELINL
CHLORIDE CHANNEL PROTEIN ; mRNA sequence.
ACCESSION A1721275
VERSION A1721275.1 GI:5038531
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 501)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (infoimage.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from Gibco
High quality sequence stop: 394.
FEATURES
Source
1..501
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:235263"
/clone_1lb="Barstead colon HPLR87"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/note="Organ: colon; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site: 1: EcoRI; Site: 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCATCTGAGTGGAGCGGCGCTACTCTTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTACTAGTAT 3' and 5' AATTACTAGT 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library constructed by Bob Barstead."
BASE COUNT 158 a 83 c 106 g 154 t
ORIGIN
Query Match 28.5%; Score 480; DB 10; Length 501;
Best Local Similarity 98.0%; Pred. No. 8.4e-113;
Matches 486; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1079 ATTAATCTGACTGAGACGCTCTGGGAGATTAATGACATGAGACGCTCACAAGTAT 1088
|||||
Db 501 ATTAATCTGACTGAGACGCTCTGGGAGATTAATGACATGAGACGCTCACAAGTAT 442
QY 1089 ATCAATTCGAATTAAGTACAAATATCTGATCTCGAGACAAATGTCATTAATCTCTTCA 1148
|||||

Db	441	ATCATGTGGAAATGAAGTACAAAGTATTTCTTGATCTCAGAGACAAAGTCAATGAATCTCTTCAA	382
Oy	1149	GTGAATACTACTGTCTGTCTCATCCCAAAAGGAAGCCAACTCTGAGGAAGTCTTTTGTTTAAA	1208
Db	381	GTGATATCTACCGCTGTCAATCCCAAAAGGAAGCCAACTCTGAGGAAGTCTTTTGTTTAAA	322
Oy	1209	CCAAATAAATTAATCTTTTGAAAATGGCAGACAGATCTTTTATTTGCTATTCAGGCTTTGAT	1268
Db	321	CCAAATAAATTAATCTTTTGAAAATGGCAGACAGATCTTTTATTTGCTATTCAGGCTTTGAT	262
Oy	1269	AAGTGCATCTGAAATCAGAAATATCCAAATTCACAGAGTATCTTTGTTTATTTCTCTCA	1328
Db	261	AAGTGCATCTGAAATCAGAAATATCCAAATTCACAGAGTATCTTTGTTTATTTCTCTCA	202
Oy	1329	CAGACTCCGCGCAGAGACACTAGTCCGTGATGAAAGCTGCGCTCCCTGCTAAATTTCAAT	1388
Db	201	CAGACTCCGCGCAGAGACACTAGTCCGTGATGAAAGCTGCGCTCCCTGCTAAATTTCAAT	142
Oy	1389	ATCAACAGCAGACATCTCTGGCATTTCAATTTTAAAAATTTATGTGNAAGTGAATAGAGAA	1448
Db	141	ATCAACAGCAGACATCTCTGGCATTTCAATTTTAAAAATTTATGTGGAAGTGAATAGAGAA	82
Oy	1449	CTGCAGCTGTCAATAGCCTTAGGGCTGAATTTTGTGAGATAAATAAATAATTAATTCATTAT	1508
Db	81	CTGCAGCTGTCAATAGCCTTAGGGCTCATTTTGTGCTATTAATAACAAAATAATCATTCAT	22
Oy	1509	CCTTTTTTTTGATTTAT	1524
Db	21	CCTTTTTTTTATTTAT	6

```

RESULT 14
AM361521
LOCUS
DEFINITION QV2-CT0261-261099-011-f03 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361521
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
REFERENCE
AUTHORS
TITLE
COMMENT
1 (bases 1 to 657)
HCGP http://www.ludwig.org.br/RESFES.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=QV2&t2=QV2-CT0261)
261099-011-f03&t3=1999-10-26&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 133
High quality sequence stop: 656.
Location/Qualifiers
1..657

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CR0261"
/clone_stage="Adult"
/notes="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESF5 PCR (U.S. Letters Patent application No. 1995-
0716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low

```

BASE COUNT	175 a	160 c	164 g	158 t
ORIGIN	stringency conditions."			
Query=Watch	28.0%	Score 470.6;	DB 10;	Length 657;
Best Local Similarity	98.1%;	Pred. No. 2.3e-110;		
Matches 476;	Conservative 0;	Mismatches 9;	Indels 0;	Gaps 0;
CY	1	AACAAAGTGGGCGCCATCATCCACACAGTCGGCTTTGGGGGCCCTCTGCACCTCAAGAATCAG	60	
DB	13	AACAAGTGGTGTCCTCTATCCACACAGTCGCTTTGGGGGCCCTCTGCACCTCAAGAATCAG	232	
OY	61	AGGACCTCTCCAAAATGACAGAGGTTTACACACATATGCTTCAGATCAAGTTCAGAAC	120	
DB	233	AGGACCTCTCCAAAATGACAGAGGTTTACACACATATGCTTCAGATCAAGTTCAGAAC	292	
OY	131	ATGGCTCATTTATGCTTTTGGGGCCCTTTTCATCAGGAATGGACTGTCTCTACGCCCT	180	
DB	233	ATGGCTCATTTATGCTTTTGGGGCCCTTTTCATCAGGAATGGACTGTCTCTACGCCCT	352	
OY	131	CCATCCACCTTGAGAGTAAGGATTAAACCCTCCACAAACAGCACAGTAGAATGACACAG	240	
DB	333	CCATCCACCTTGAGAGTAAGGATTAAACCCTCCACAAACAGCACAGTAGAATGACACAG	412	
OY	211	TGATCTGTGACAGACACCGTGGGAAGAAGACACTTTTCTTATCACTCTGACACACGAC	300	
DB	413	TGATCTGTGACAGACACCGTGGGAAGAAGACACTTTTCTTATCACTCTGACACACGAC	472	
OY	301	CTCCCCAAATCCTTCTCGGGATGCCAGTAGGACACAGAAAGTAGGCTTTGATGAGACA	360	
DB	473	CTCCCCAAATCCTTCTCGGGATGCCAGTAGGACACAGAAAGTAGGCTTTGATGAGACA	532	
CY	311	AAAAACACCAAATGGCCCTACCTCCAAATCCAGGCAATTCCTAAGTTGGCACTTGGAAAT	420	
DB	513	AAAAACACCAAATGGCCCTACCTCCAAATCCAGGCAATTCCTAAGTTGGCACTTGGAAAT	592	
OY	421	ACAGCTCTGAAGCAAGCTCACAACACTTGACCTCAGCTCAGCCGCTGCTCATAAG	480	
DB	513	ACAGCTCTGAAGCAAGCTCACAACACTTGACCTCAGCTCAGCTCAGCCGCTCATAAG	652	
OY	411	CTACCC 405		
DB	613	CTACCC 657		
RESULT	15			
LOCUS	AM361520	621 bp	mRNA	EST
DEFINITION	QV2-CT0261-261099-011-e07 CT0261 Homo sapiens CDNA, mRNA sequence.			
ACCESSION	QV2-CT0261-261099-011-e07			
VERSION	AM361520.1	GI:8666274		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	HCGP http://www.ludwig.org.br/ORSTES.			
JOURNAL	The FAPESP/LICR Human Cancer Genome Project			
COMMENT	Unpublished (1999) Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil. Tel.: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be found in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?rl-QV2&t2-QV2-CT0261- 261099-011-e07&t3-1999-10-26&t4-1)			

Seq primer: puc 18 forward
High quality sequence start: 88
High quality sequence stop: 613.
Location/Qualifiers

FEATURES

1..621

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="CT0261"

/dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:

SmaI: A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196

,716 - Ludwig Institute for Cancer Research) profiles

into the puc 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

BASE COUNT 166 a 151 c 150 g 154 t

ORIGIN

Query Match 27.7%; Score 465.4; DB 10; Length 621;

Best Local Similarity 97.6%; Pred. No. 4.9e-109;

Matches 483; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

```
OY 1 AACAAAGTGTGTCATCCATCCACAGTGTGTTGGGGCCCTCTGAGCTCAGAACTAG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 AACATAGTGTGTCCTCCATCCACAGTGTGTTGGGGCCCTCTGAGCTCAGAACTAG 183
OY 61 AG-GAGCTGTCCAAATGACAGAGTTACAGACATATGCTTCAGATCAAGTTGAGAC 119
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 AGTGAAGCTGTCCAAATGACAGAGTTACAGACATATGCTTCAGATCAAGTTGAGAC 243
OY 120 AATGACCTCATATGATGCTTTGGGGCCCTTTCATCAGAAATGAGCTCTCTCAGCCG 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 AATGGCTCATATGATGCTTTGGGGCCCTTTCATCAGAAATGAGCTCTCTCAGCCG 303
OY 180 TCCATCCAGCTTGAAGATAGGATTACCTCCAGACAGCCAGTGATGAATGGCACA 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 TCCATCCAGCTTGAAGATAGGATTACCTCCAGACAGCCAGTGATGAATGGCACA 363
OY 240 GTGATCGTGGACAGCAGCGTGGAAAGACACTTGTCTTATCAGCTGGAACAAGCAG 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 364 GTGATCGTGGACAGCAGCGTGGAAAGACACTTGTCTTATCAGCTGGAACAATGCA 423
OY 300 CCTGCCCAATCTCTCTGGGATCCCACTGGACAGAAAGCAAGTGGCTTTGATGAGAC 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 CCTGCCCAATCTCTCTGGGATCCCACTGGACAGAAAGCAAGTGGCTTTGATGAGAC 483
OY 360 AAAAAGACCAAAATGCGCTACTCTCAATCCAGGCAATGCTTAAGTTGGCACTGGAAA 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 AAAAAGACCAAAATGCGCTACTCTCAATCCAGGCAATGCTTAAGTTGGCACTGGAAA 543
OY 420 TACAGTGTGCAAGAGTCAAAACCTTGACCTGACGTACGTCGCCGTGGTCCAT 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 TACAGTGTGCAAGAGTCAAAACCTTGACCTGACGTACGTCGCCGTGGTCCAT 603
OY 480 GCTACCTGCTTCCA 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 604 GCTACCTGCTTCCA 618
```

Search completed: April 3, 2002, 20:28:45
Job time: 40288 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:55:58 ; Search time 521.98 Seconds
(without alignments)
730.224 Million cell updates/sec

Title: US-09-049-696-19

Perfect score: 1683
Sequence: 1 AACAACTGGTGCATCATC.....AAATGCTAACACACTGGGTA 1683

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCBUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	790.8	47.0	878	1	US-08-469-667-8
2	790.8	47.0	878	5	PCR-US95-07289-8
3	228.2	13.6	576	4	US-09-385-982-23
4	221.4	13.2	595	4	US-09-385-982-25
5	200.8	11.9	618	4	US-09-385-982-24
6	183.4	10.9	611	4	US-09-385-982-27
7	168.6	10.0	742	4	US-09-385-982-33
8	41.4	2.5	7218	1	US-08-392-625-16
9	36.4	2.2	8700	2	US-08-466-961A-16
10	36.4	2.2	8700	2	US-08-466-961A-16
11	36.4	2.2	8700	2	US-08-645-193B-18
12	36	2.1	3158	2	US-08-464-517-36
13	36	2.1	3158	2	US-08-246-361A-36
14	36	2.1	3158	2	US-08-463-772-36
15	35.2	2.1	3058	4	US-09-276-531-107
16	34.8	2.1	805	1	US-08-118-469A-6
17	34.8	2.1	805	1	US-08-909-119-6
18	34.8	2.1	19124	2	US-08-487-826B-13
19	34.6	2.1	5923	4	US-09-064-922-3
20	33.8	2.0	1566	4	US-08-871-572B-2
21	33.8	2.0	2255	4	US-08-871-572B-3
22	33.6	2.0	3183	1	US-08-413-135-1
23	33.6	2.0	4090	4	US-08-569-214-4
24	33.6	2.0	4090	4	US-08-937-236-4
25	33.6	2.0	4810	3	US-08-852-629-11
26	33.2	2.0	2520	2	US-08-454-557C-50
27	33.2	2.0	2520	2	US-08-340-426D-50

28	33.2	2.0	2520	2	US-08-450-673C-50	Sequence 50, Appl
29	33.2	2.0	2520	5	PCR-US95-1711A-50	Sequence 50, Appl
30	33.2	2.0	3095	6	5231168-1	Patent No. 5231168
31	33.2	2.0	10684	3	US-08-618-100B-3	Sequence 3, Appl
32	33.2	2.0	1984	1	US-07-885-970A-25	Sequence 25, Appl
33	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
34	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
35	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
36	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
37	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
38	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
39	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
40	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
41	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
42	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
43	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
44	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl
45	33	2.0	1985	1	US-08-298-687A-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-08-469-667-8
Sequence 8, Application US/08469667
Patent No. 5733748
GENERAL INFORMATION:
APPLICANT: YU, Guo-liang
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gillfillan, Cecchi,
STREET: Stewart & Olstein
CITY: Roseland
STATE: NJ
COUNTRY: USA
FID: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,667
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Petrarro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-435
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1744
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
US-08-469-667-8

QueryMatch 47.0%; Score 790.8; DB 1; Length 878;
Best local Similarity 97.9%; Pred. No. 2.3e-230;
Matches 820; Conservative 1; Mismatches 15; Indels 2; Gaps 2;


```

QY 692 TGTCTACTCAAGGATATTTACAACTTATGACACCAAGTGTAGATACAGTGTAAAAGTCCG 751
  |||
  |||
  |||
Db 1 TGTCTACTCAAGGATATTTACAACTTATGACACCAAGTGTAGATACAGTGTAAAAGTCCG 60
  |||
  |||
  |||
QY 752 GCGCTCTGGAGAGATTAACGACGACGAGAGATGTATACCCAGCAGATGGAGACT 811
  |||
  |||
  |||
Db 61 GCGCTCTGGAGAGATTAACGACGACGAGAGATGTATACCCAGCAGATGGAGACT 120
  |||
  |||
  |||
QY 812 GTACATACCTGGCTGGATTTGAAATGATGAATACAAATGAAATCCACCAAGACCTGAAT 871
  |||
  |||
  |||
Db 121 GTACATACCTGGCTGGATTTGAAATGATGAATACAAATGAAATCCACCAAGACCTGAAT 180
  |||
  |||
  |||
QY 872 TAATAAGAGATGATTTCAACAACAAGAGTGTGTTTCAGCAACAATCCTCGGAGGCTC 931
  |||
  |||
  |||
Db 181 TAATAAGAGATGATTTCAACAACAAGAGTGTGTTTCAGCAACAATCCTCGGAGGCTC 240
  |||
  |||
  |||
QY 932 ATTGTGCTCTGATGTCGCCAAATGCTCCATACCTGATCTCTCCACCTGGCCAAAT 991
  |||
  |||
  |||
Db 241 ATTGTGCTCTGATGTCGCCAAATGCTCCATACCTGATCTCTCCACCTGGCCAAAT 300
  |||
  |||
  |||
QY 992 CACGACCTGGAAGGCGGAAATTCACGGGGGAGTCTCATTAATCTGACTTGGACAGCTCC 1051
  |||
  |||
  |||
Db 301 CACGACCTGGAAGGCGGAAATTCACGGGGGAGTCTCATTAATCTGACTTGGACAGCTCC 360
  |||
  |||
  |||
QY 1052 TGGGAGATGATTAAGACCAAGAGTGTGTTTCAGCAACAATCCTCGGAGGCTC 1111
  |||
  |||
  |||
Db 361 TGGGAGATGATTAAGACCAAGAGTGTGTTTCAGCAACAATCCTCGGAGGCTC 420
  |||
  |||
  |||
QY 1112 TCTGATCTGAGAGCAATGATCAATGATCTCTCAATGATTAATCTGCTCATCC 1171
  |||
  |||
  |||
Db 421 TCTGATCTGAGAGCAATGATCAATGATCTCTCAATGATTAATCTGCTCATCC 480
  |||
  |||
  |||
QY 1172 AAAGAACCAACTCTGAGAACTTTTGTGTTTAAACCAAGAAACATTAATTTGAAA 1231
  |||
  |||
  |||
Db 481 AAAGAACCAACTCTGAGAACTTTTGTGTTTAAACCAAGAAACATTAATTTGAAA 540
  |||
  |||
  |||
QY 1232 TGGCAGAGATCTTTTCATGCTGATTCAGGCTGTTGATTAAGTCTGCTGAAATCA 1291
  |||
  |||
  |||
Db 541 TGGCAGAGATCTTTTCATGCTGATTCAGGCTGTTGATTAAGTCTGCTGAAATCA 600
  |||
  |||
  |||
QY 1292 ATCCAACTGAGAGATCTTTTGTGTTTAACTCCACAGACTCCGCCAGAGACCTAG 1351
  |||
  |||
  |||
Db 601 ATCCAACTGAGAGATCTTTTGTGTTTAACTCCACAGACTCCGCCAGAGACCTAG 660
  |||
  |||
  |||
QY 1352 TCTGATGAAAGCTGCTCTCTGT-TCTAATATTCATATCAACAGCAATCTCTGCA 1410
  |||
  |||
  |||
Db 661 TCTGATGAAAGCTGCTCTCTGT-TCTAATATTCATATCAACAGCAATCTCTGCA 720
  |||
  |||
  |||
QY 1411 TTCACATTTTAAAAATTTATGGAAGTGGATAGGAACTGCACTGCAATAGCTTAG 1470
  |||
  |||
  |||
Db 721 TTCACATTTTAAAAATTTATGGAAGTGGATAGGAACTGCACTGCAATAGCTTAG 780
  |||
  |||
  |||
QY 1471 GCTGAATTTTGTGATTAATTAATTAATCAATCAATCTTTTGTGATTAATAA 1528
  |||
  |||
  |||
Db 781 GCTGAATTTTGTGATTAATTAATTAATCAATCAATCTTTTGTGATTAATAA 837
  |||
  |||
  |||

```

```

RESULT 2
PCT-US95-07289-8
: Sequence 8, Application PC/TUS9507289
: GENERAL INFORMATION:
: APPLICANT: Yu, Guo-Liang
: APPLICANT: Rosen, Craig
: TITLE OF INVENTION: Colon Specific Genes and Proteins
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carella, Byrne, Bain, Giffillan, Cecchi,
: ADDRESS: Stewart & Olstein
: STREET: 6 Becker Farm Road
: CITY: Roseland
: STATE: NJ
: COUNTRY: USA
: ZIP: 07068-1739

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: PCT/US95/07289
  FILING DATE: 06-JUN-1995
  CLASSIFICATION:
  ATORNEY/AGENT INFORMATION:
  NAME: Ferraro, Gregory D.
  REGISTRATION NUMBER: 36,134
  REFERENCE/DOCKET NUMBER: 325800-265
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 201-994-1700
  TELEFAX: 201-994-1744
  INFORMATION FOR SEQ ID NO: 8:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 878 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  FEATURE:
  NAME/KEY: CDS
  LOCATION: 2..685
PCT-US95-07289-8

```

```

Query Match 47.0%; Score 790.8; DB 5; Length 878;
Best local Similarity 97.9%; Pred. No. 2.3e-210;
Matches 820; Conservative 1; Mismatches 15; Indels 2; Gaps 2;

QY 632 TGTCTACTCAAGGATATTTACAACTTATGACACCAAGTGTAGATACAGTGTAAAAGTCCG 751
  |||
  |||
  |||
Db 1 TGTCTACTCAAGGATATTTACAACTTATGACACCAAGTGTAGATACAGTGTAAAAGTCCG 60
  |||
  |||
  |||
QY 752 GCGCTCTGGAGAGATTAACGACGACGAGAGATGTATACCCAGCAGATGGAGACT 811
  |||
  |||
  |||
Db 61 GCGCTCTGGAGAGATTAACGACGACGAGAGATGTATACCCAGCAGATGGAGACT 120
  |||
  |||
  |||
QY 812 GTACATACCTGGCTGGATTTGAAATGATGAATACAAATGAAATCCACCAAGACCTGAAT 871
  |||
  |||
  |||
Db 121 GTACATACCTGGCTGGATTTGAAATGATGAATACAAATGAAATCCACCAAGACCTGAAT 180
  |||
  |||
  |||
QY 872 TAATAAGAGATGATTTCAACAACAAGAGTGTGTTTCAGCAACAATCCTCGGAGGCTC 931
  |||
  |||
  |||
Db 181 TAATAAGAGATGATTTCAACAACAAGAGTGTGTTTCAGCAACAATCCTCGGAGGCTC 240
  |||
  |||
  |||
QY 932 ATTGTGCTCTGATGTCGCCAAATGCTCCATACCTGATCTCTCCACCTGGCCAAAT 991
  |||
  |||
  |||
Db 241 ATTGTGCTCTGATGTCGCCAAATGCTCCATACCTGATCTCTCCACCTGGCCAAAT 300
  |||
  |||
  |||
QY 992 CACGACCTGGAAGGCGGAAATTCACGGGGGAGTCTCATTAATCTGACTTGGACAGCTCC 1051
  |||
  |||
  |||
Db 301 CACGACCTGGAAGGCGGAAATTCACGGGGGAGTCTCATTAATCTGACTTGGACAGCTCC 360
  |||
  |||
  |||
QY 1052 TGGGAGATGATTAAGACCAAGAGTGTGTTTCAGCAACAATCCTCGGAGGCTC 1111
  |||
  |||
  |||
Db 361 TGGGAGATGATTAAGACCAAGAGTGTGTTTCAGCAACAATCCTCGGAGGCTC 420
  |||
  |||
  |||
QY 1112 TCTGATCTGAGAGCAATGATCAATGATCTCTCAATGATTAATCTGCTCATCC 1171
  |||
  |||
  |||
Db 421 TCTGATCTGAGAGCAATGATCAATGATCTCTCAATGATTAATCTGCTCATCC 480
  |||
  |||
  |||
QY 1172 AAAGAACCAACTCTGAGAACTTTTGTGTTTAAACCAAGAAACATTAATTTGAAA 1231
  |||
  |||
  |||
Db 481 AAAGAACCAACTCTGAGAACTTTTGTGTTTAAACCAAGAAACATTAATTTGAAA 540
  |||
  |||
  |||
QY 1232 TGGCAGAGATCTTTTCATGCTGATTCAGGCTGTTGATTAAGTCTGCTGAAATCA 1291
  |||
  |||
  |||
Db 541 TGGCAGAGATCTTTTCATGCTGATTCAGGCTGTTGATTAAGTCTGCTGAAATCA 600
  |||
  |||
  |||

```


QY 1292 ATCCAGATTCGACAGATTCCTTTTCTCCACAGACTCCGCCAGACACCTAG 1351
DB 601 ATCCAGATTCGACAGATTCCTTTTCTCCACAGACTCCGCCAGACACCTAG 660
QY 1352 TCCGATGGAAGCTGCTGCTCTGT-CCTAATTCATATCAACACACACCTCTCGCA 1410
DB 661 TCCGATGGAAGCTGCTGCTCTGT-CCTAATTCATATCAACACACACCTCTCGCA 720
QY 1411 TTCACATTTTAAATTAATTCAGATTCGATGAGAGAACTGACCTGCAATGACCTAG 1470
DB 721 TTCACATTTTAAATTAATTCAGATTCGATGAGAGAACTGACCTGCAATGACCTAG 780
QY 1471 GCTGATTTTGTGCATATAATAATAATCAATTCATCTTTTGTGATATAAA 1528
DB 781 GGTGATTTTGTGCATATAATAATAATCAATTCATCTTTTGTGATATAAA 837

RESULT 3

US-09-385-982-23
Sequence 23, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCDA-260XX
CURRENT FILING DATE: 1999-08-30
EARLIER FILING DATE: 1999-06-08
EARLIER FILING DATE: 1999-01-27
EARLIER FILING DATE: 1999-01-27
EARLIER FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(576)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-23

Query Match 13.6%; Score 228.2; DB 4; Length 576;
Best Local Similarity 69.1%; Pred. No. 1.3e-59;
Matches 385; Conservative 0; Mismatches 163; Indels 9; Gaps 5;

QY 582 CCAATTCAGGGCCAGCTGCACAGCCCGATGATCACTGATGGAAGAAACAGTTACC 641
DB 2 CCGCTCTCTGGAGCCCAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 61
QY 642 TTGGAAGTACTGATGATGAGAGAGGCTGCTGATGATGATGATGATGATGATGATGAT 701
DB 62 TTGGAAGTACTGATGATGAGAGAGGCTGCTGATGATGATGATGATGATGATGATGAT 121
QY 702 AGGTATTTACACATTTATACACGAATGCTGATGATGATGATGATGATGATGATGAT 761
DB 122 AGGTATTTACACATTTATACACGAATGCTGATGATGATGATGATGATGATGATGAT 181
QY 762 GGAGTTTACAGCAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 821
DB 182 GGAGTTTACAGCAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 241
QY 822 GGCTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 881
DB 242 GGCTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 298
QY 882 GATGTTTCAACAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 941

DB 299 GATACCAAGCACCCTGGAGATTTCAAGCAGACAGATCCGAGATGATGATGATGAT 358
QY 912 TCGATGATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001
DB 912 TCGATGATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 417
QY 1002 TCGATGATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1061
DB 1002 TCGATGATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 474
QY 1062 TCGATGATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1121
DB 1062 TCGATGATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 532
QY 1122 AGAGACAGTTCAATGA 1138
DB 1122 AGAGACAGTTCAATGA 549

RESULT 4

US-09-385-982-25
Sequence 25, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCDA-260XX
CURRENT FILING DATE: 1999-08-30
EARLIER FILING DATE: 1999-06-08
EARLIER FILING DATE: 1999-01-27
EARLIER FILING DATE: 1999-01-27
EARLIER FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 25
LENGTH: 595
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(595)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-25

Query Match 13.2%; Score 221.4; DB 4; Length 595;
Best Local Similarity 68.4%; Pred. No. 1.5e-57;
Matches 355; Conservative 0; Mismatches 155; Indels 9; Gaps 4;

QY 844 ACATACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
DB 844 ACATACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 873 ACATACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933
DB 873 ACATACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 117
QY 933 ACATACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 177
DB 933 ACATACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 234
QY 1034 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1113
DB 1034 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 294

OY	1114	TTGATTCAGAGAACATGTCGAATCTCTCAATGTAATCTATGCTGTCATCCGA	1173
Db	295	ttgatccaaggaacagcttgcgtgacgcttccaagyaactactactgctcgtaccaa	354
OY	1174	AGGAAGCCAACTCTGACGAAGTCTTTTGTGTTTAAACACAAAAACAATTACTTTTGAAATG	1233
Db	355	agggagcgcaactccaangaaagcttgcgtttaaaccaggaaataatctcagaagaatg	414
OY	1234	GCACAGATCTTTTCATTGCTAATTCACGGCTGTGATTAAGTCGATCTGGAATCGAAGAATAT	1293
Db	415	caaccacacatattatgcttccttcaaaagatagatanaa--caatttgcac--naagt	471
OY	1294	CCAACTATGCGACGATCTTTCTGTTAATTCCTCCACAGA	1332
Db	472	tccaatctgmaccagtaactctggttatccctccacga	510

```

RESULT      5
US-09-385-982-24/C
: Sequence 24, Application US/09385982
: Patent No. 6262334
: GENERAL INFORMATION:
: APPLICANT: ENDEGE, WILSON O., ET AL.
: TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
: TITLE OF INVENTION: PRODUCTS: II
: FILE REFERENCE: CCDNA-260XX
: CURRENT APPLICATION NUMBER: US/09/385,982
: EARLIER APPLICATION NUMBER: 09/328,111
: EARLIER FILING DATE: 1999-06-08
: EARLIER APPLICATION NUMBER: 60/117,393
: EARLIER FILING DATE: 1999-01-27
: EARLIER APPLICATION NUMBER: 60/098,639
: EARLIER FILING DATE: 1998-08-31
: NUMBER OF SEQ. ID NOS.: 544
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 24
: LENGTH: 618
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(618)
: OTHER INFORMATION: n = A,T,C or G
: US-09-385-982-24

```

Query Match	11.9%	Score 200.8	DB 4	Length 618
Best Local Similarity	69.5%	Pred No. 2.8e-51		
Matches 303	Conservative 0	Mismatches 127	Indels 6	Gaps
QY 951	CCAAATGCTCCCATACCTGATCTCTCCACCTGGCCAAATCACCAGCCTGAGGCGGAA	1010		
Db 516	CCCAAGCCTTCCTTGGTCTGTCGCAATACCCACCAAGCAATACAGACCTTGATCCACA	457		
QY 1011	ATTACGGGGGACGTGTCATTATCGACTTGGAGACGCTCCGGGAGATTAATGACCAT	1070		
Db 456	GTTCATATAG---ATAGATTATCTTCAATGACACGCACCGAGATTAATTTGATGTT	400		
QY 1071	GGAACAGCTCACAAGTATATCATTCGAAATAGTACAAAGTATTTCTGATCTCAGAGACAG	1130		
Db 399	GGAAGATTCACGTTATATATCATTAAGAAATAGTCAGATTTCTGATCTAAGACAGT	340		
QY 1131	TTTCAATAACTCTTCAAGTAACTACTACTGCTGTCATCCCAAGAAAGCAACCTCTAG	1190		
Db 339	TTTGATGATGCTCTTCAAGTAAATACTACTGATGCTGTCACCAAGAGAGCCCAACTCCAG	280		
QY 1191	GAACTCTTTTGTGTTAAACCGAAACCATTTACTTTTGAAGAAATGGACAGATCTTTCAT	1250		
Db 279	GAAAGCTTTTCATTTTAAACCGAAATATCTCAGAAAGAAATGCAACCCACATTTATT	220		
QY 1251	GCATTCAGGCTGTGTATTAAGGTCGATCTGGAATCAGAAATATCCAAATTCGACATTCAGAGTA	1310		

Db	219	GCATTTAAAGTATAGATATAAAGCAATTTCACATCAAAATATTCACCAATTCGCACAAAGTA	160
Gy	1311	TCTTTGTTTATTCCTCCACAGACATCTGCGCCGACAGA ---CACTACTCTGATATGAAGCTCT	136
Db	139	ACTTTGTTTATTCCTCCACAGCAATTCCTGATGACATTGATCTTCACCTACTCTCTCTCT	100
Gy	1368	GCTCTCTTCTCTTAATA 1383	
Db	99	ACTCTTACTCTTGATA 84	

```

RESULT 6
US-09-385-982-27
Sequence 27, Application US/09385982
Patent No. 626234
GENERAL INFORMATION
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS: II
FILE REFERENCE: CCDNA-2603X
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 344
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 27
LENGTH: 611
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(611)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-27

```

Query Match	10.9%	Score 183.4	DB 4	Length 611
Best local Similarity	68.0%	Pred No. 5,3e-46		
Matches 300	Conservative 0	Mismatches 136	Indels 5	Gaps
QY 512	CCAAATTCAGGCGCAGTGTACAGCCGCTATTTGAATCACTGAAT-GGAAAAACACTTAC	640		
DY 1				
DY 2	cctgtctcttgagacgaatgactgactgtcttcatcttgatcacaagatgggacatacagaagt	61		
QY 641	CTTGGAACCTACTGGATATATGGAGCAGGCTGCTGATTTGCTACTTAAGATGACGGTGTACTTC	700		
DY 1				
DY 52	cttggaaacttttgataatggttcagcgctgacttcttcaaaagatgtaagatcttacc	121		
QY 701	AGGATATTTCAACAATCTATACACAGATGTAATATACGTGTAAATGGTGGGGCTGTGGG	760		
DY 1				
DY 122	caggtatttttaacgcataatacagaataatggcagatataagcttaaaagtttcggctatgg	181		
QY 761	AGGAGTTTAAACCGACGCCAGACGGAGAGTGTATACCCACAGACAGTGGAGACCTGTATACC	820		
DY 1				
DY 132	agagacaaacacttgcaggtctaaatctaacgcttccacttgaatagaagccgcyttacatcc	241		
QY 821	TGGTGATTTAGATATATCAATATACATATGGATTCACCAAGCTCGAATTTATATGAAGA	880		
DY 1				
DY 242	aagcttgytctgtgaagaagggaataatgaagcaaaccccgcaagccggaatctgat---ga	298		
QY 881	TGATGTTTAAACACAGCAAGCAAGTGTGTTTTCACAGACACATCTCTGGAGGCTCAATTTGTGC	940		
DY 1				
DY 209	ggatactcagacacaccttggagatcttcagccgaacagatcccgaggytgcatcttgggt	358		
QY 941	TTTC-TGATGTCCCAAAATGCTCCCATACCTCAATCTCTTCCCACTGTGCCCAATATACCGACAC	999		
DY 1				
DY 359	atcccaaatgctccaaaccttctgcctgtaaccaataaccaccaagtaataatcaacagacc	418		

Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:21:09 ; Search time 1321.64 Seconds

(without alignments)
1091.733 Million cell updates/sec

Title: US-09-049-696-19

Perfect score: 1683
Sequence: 1 AACCAAGCTGTCATCANC.....AAATGCTAACCACTGGCTA 1683

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101:*

- 1: /SIDSL/gcgdata/geneseq/geneseqn/NA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqn/NA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqn/NA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqn/NA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqn/NA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqn/NA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqn/NA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqn/NA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqn/NA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqn/NA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqn/NA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqn/NA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqn/NA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqn/NA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqn/NA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqn/NA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqn/NA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqn/NA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqn/NA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqn/NA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqn/NA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683	100.0	3109	22	AAH35019 Human colon cancer
2	1683	100.0	3111	20	AAZ09840 Human membrane spa
3	1671	99.3	3311	22	AAI29502 C902P determined c
4	1528.4	90.8	2854	22	AAH34879 Human colon cancer
5	1528.4	90.8	2854	22	AAH34879 Human colon cancer
6	1526.8	90.7	2867	22	AAH33285 Human secreted pro
7	1517.4	90.2	2825	22	AAH46124 Human ICAC-1 CDNA,
8	1469	87.3	2745	20	AAH46124 Human ICAC-1 nucl
9	1466	87.1	2742	22	AAH46102 Human ICAC1 coding
10	940.2	55.9	2843	22	AAH46120 Mouse Gob-5 CDNA,
11	939.4	55.8	2931	20	AAH81925 Murine ICAC-1 nuc

12	911.6	54.2	2739	22	AAH46101	Mouse Gob-5 coding
13	790.8	47.0	878	18	AAH45884	Human colon specific
14	790.8	47.0	878	18	AAH45884	Polynucleotide seq
15	667.6	39.7	2616	21	AAH46435	Clone 2516888 of a
16	667.6	39.7	3265	21	AAH46435	Membrane-bound pro
17	667.6	39.7	3265	22	AAH46435	Human PRO1124 CDNA
18	667.6	39.7	3265	22	AAH46435	Human colon cancer
19	654	38.9	1802	22	AAH33192	Human PRO1124 (UNQ
20	654	38.9	1802	22	AAH33192	Human colon cancer
21	326	19.4	486	22	AAH46121	Mouse Gob-5 CDNA
22	304	18.1	2784	20	AAH46121	Human lung tumor a
23	304	18.1	2784	21	AAH46121	Human lung cancer-
24	301.6	17.9	3156	21	AAH46121	Human membrane Cha
25	301.6	17.9	3951	20	AAH46121	Human lung tumor a
26	301.6	17.9	3951	21	AAH46121	Human lung cancer-
27	301.6	17.9	8031	21	AAH46121	Human lung cancer-
28	298.4	17.7	2454	22	AAH46121	Human tumor-associ
29	279.6	16.6	3190	20	AAH46121	Human ICAC-2 nucl
30	241.2	14.3	255	16	AAH46121	Human gene signatu
31	228.2	13.6	576	21	AAH46121	Human colon cancer
32	221.4	13.2	595	21	AAH46121	Human colon cancer
33	200.8	11.9	618	21	AAH46121	Human colon cancer
34	183.4	10.9	611	21	AAH46121	Human colon cancer
35	168.6	10.0	742	21	AAH46121	Human colon cancer
36	124.2	7.4	546	20	AAH46121	Human lung tumor a
37	124.2	7.4	546	21	AAH46121	Human lung cancer-
38	115.4	6.9	585	20	AAH46121	Human lung tumor a
39	115.4	6.9	590	21	AAH46121	Human lung cancer-
40	115.4	6.9	3362	21	AAH46121	Human lung cancer-
41	113.8	6.8	3362	20	AAH46121	Human lung tumor a
42	113.6	6.7	936	22	AAH46121	Oligonucleotide D1
43	113.6	6.7	936	22	AAH46121	Oligonucleotide D1
44	113.6	6.7	936	22	AAH46121	Oligonucleotide D1
45	113.6	6.7	936	22	AAH46121	Oligonucleotide D2

ALIGNMENTS

AAH35019	standard; cDNA; 3109 BP.
AAH35019	
03-SEP-2001	(first entry)
Human colon cancer antigen encoding CDNA SEQ ID NO:2101.	
Human colon cancer; colon cancer antigen; diagnosis; detection;	
colorectal carcinoma; ss.	
Human sapiens.	
WG200122920-A2.	
05-APR-2001.	
28-SEP-2000; 2000MO-US26524.	
29-SEP-1999; 99US-0157137.	
03-NOV-1999; 99US-0163280.	
(HUMA-) HUMAN GENOME SCI INC.	
Ruben SM, Barash SC, Birse CE, Rosen CA;	
WRI; 2001-235357/24.	
P-FSDB; AAG75614.	
Nucleic acids encoding 4277 human colon cancer-associated polypeptides,	
useful for preventing, diagnosing and/or treating colorectal cancers -	

PS Claim 1: Page 3587-3588; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing P.
CC Inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

CC Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other:

Query Match 100.0%; Score 1683; DB 22; Length 3109;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1683; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAAAGTGTGTCATCCACAGAGTGGGCGCTTTCAGTCAAGAACTAG 60
DB 1166 aacaaagtggtgtcattccatccacacagtcgttggtggtcgtcgaagactag 1223
QY 61 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATGCTTCAATCAAGTTCAGACA 120
DB 1224 aggaagctgtccaaatgacagaggtttacagacatgcttcaatcaagttcagaaca 1283
QY 121 AAGGCTCTTGTATGCTTTGGGCGCTTTCAGTCAAGAACTAGTCTCTCAGCGCT 180
DB 1284 atggctcttgtatgctttgggcgcttttcagtcaagaaactagctctcagcgct 1343
QY 181 CCATCAGCTTGAGAGTAAGGATTAACCTCCAGAACGCCAGTGAATGACAGC 240
DB 1344 ccattcagcttgagagtaaggatttaacctccagaaagcagtgagatgagatgagacag 1403
QY 241 TGATCTGAGACAGACCGTGGGAAAGAGACATTTGTTATACCTGAGACAGCAGC 300
DB 1404 tgatctgagacagacggtgggaaagagacatttgttatccctgagacagacagcagc 1463
QY 301 CTCGCCAAATCTTCTCTGGGATCCAGTGCAGAGAGAGAGGAGGCTTGTAGTGACA 360
DB 1464 ctccgcaaatctctctctggtatccagtgagacagagagagtggtctttagtgagaca 1523
QY 361 AAAACACCAAAATGCGCTACCTCCAAATCCAGAGCATTTGCTAAGTGGACACTGGAAT 420
DB 1524 aaacacccaaaatgcgctacctccaaatccagagcatgtctaagttgagcaactggaat 1583
QY 421 ACAGTGTGAAGAGAGTCAAACTTGACCTGACCTGACAGTCCGCTGCTCCAAATG 480
DB 1584 acagtgtgaagagagtcaaaacttgacctgacctgacctgacctgacctgacctgacctg 1643
QY 481 CTACCTGCTCCAAATTTACAGTACTTCCAAAGAGACAGACAGACAGCAATTTCCCA 540
DB 1644 ctacctgctccaaatttacagtacttccaaagagacagacagacagacaaatctccca 1703
QY 541 GCGCTCTGAGTATTTATGCAATTTGCAAGAGAGGCTCCCAATTTAGGGCCAGT 600
DB 1704 gcgctctgagatatttatgcaattttgcaagagagagagagagagagagagagagag 1763
QY 601 TCACAGCCCTGATTTGATCAGTGAATGAGAAACAGTACTTACCTTGAATGATGAT 660
DB 1764 tcacagccctgatatttatgagatgagaaacagtgacttgaactacttgataatg 1823

QY 661 GAGCAGTGTGCTGATGCTACTAAGATGATGAGGTGTCTACTCAAGATTTACACTATG 720
DB 1824 gagcagtggtgctgactgactaagatgaggtgtctactcaagatattcacactatg 1883
QY 721 ACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 1884 acagcaatgatatgatatgatatgatatgatatgatatgatatgatatgatatgatatg 1943
QY 781 GAGAGTGTATACCCGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 1944 gagagtgatatacccgagagagtgaagagagagagagagagagagagagagagagag 2003
QY 841 AATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 2004 aatataatgatatatacccgagagagagagagagagagagagagagagagagagagag 2063
QY 901 TGCTTTACAGCAAAATCTCTCGGAGGCTTTCAGTCAAGAGAGAGAGAGAGAGAGAG 960
DB 2064 tgctttacagcaaaatctctcggagagagagagagagagagagagagagagagagag 2123
QY 961 CCATACCTGATCTCTTCCACCTGAGCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 2124 ccatacctgatctcttccacagagagagagagagagagagagagagagagagagagag 2183
QY 1021 GCAGTCTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 2184 gcagctctcatatactgactgactgactgactgactgactgactgactgactgactgact 2243
QY 1081 ACAAGTATATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 2244 acaagtataatcatattgatatgatatgatatgatatgatatgatatgatatgatatg 2303
QY 1141 CTCTTCAAGTCAATATCT 1200
DB 2304 ctcttcaagtcaatattct 2363
QY 1201 TGTTTAAACCGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
DB 2364 tgtttaaaccgaaatatttaatttaatttaatttaatttaatttaatttaatttaattta 2423
QY 1261 CTGTGTATTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 2424 ctgtgtatataagtgatgatatgatatgatatgatatgatatgatatgatatgatatg 2483
QY 1321 TTCTTCACAGAGCTCCGCGAGAGACACTTCTGATGAAAGCTCTCTCTCTCTCTCT 1380
DB 2484 ttcttcacagagctccgcgagagacacttctgataaagctctctctctctctctctct 2543
QY 1391 ATATTCAATCAACAGACACATTCCTGATGATGATGATGATGATGATGATGATGAT 1440
DB 2544 atattcaatcaacagacacattcctgataaagctctctctctctctctctctctctct 2603
QY 1441 TAGAGAACTGAGCTGATTAAGCTTGAAGCTGATGATGATGATGATGATGATGATGAT 1500
DB 2604 taggaactgagctgatataagcttgaagctgatatgatatgatatgatatgatatg 2663
QY 1501 TCATTCATCTCTTTTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
DB 2664 tcattcatctcttttgtgatttaatttaatttaatttaatttaatttaatttaattta 2723
QY 1561 GGGGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
DB 2724 ggggatataatataatataatataatataatataatataatataatataatata 2783
QY 1621 ATACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
DB 2784 atactaatatataatataatataatataatataatataatataatataatataat 2843
QY 1681 GGA 1683
DB 2844 gga 2846

OY 1441 TAGGAGAACTGACGCTGTCATATAGCTAGGCTGAATTTTTCAGATAAATAAATAA 1500
 |||||||
 Db 2750 tagagaaactgcaagctgcaataagcctagggcgaatttgcagaataaataa 2809
 OY 1501 TCATTCATCCCTTTTGTGATTAATAATTTTCAAAATGATTTTGTAGACTTCTGTAGG 1560
 |||||||
 Db 2810 tcatcattcccttttttgatataaatttctcaaaatgcatltagacttccgtag 2869
 OY 1561 GGGCGATATCTAAATGATATAGTACTTATTAATGATTAATGATTAATGATTAATGAT 1620
 |||||||
 Db 2870 gggcgataactaaatgataatgataatgataatgataatgataatgataatgataat 2929
 OY 1621 ATACGATATGATTTTGTAGCTTCTGTAGGCGGATTAATAATGATTAATAACGATGG 1680
 |||||||
 Db 2930 ataccgaatgattttagacttccgtagggcgataaataaataatgataaactgg 2989
 OY 1681 GTA 1683
 |||
 Db 2990 gta 2992
 RESULT 3
 AAI29502
 ID AAI29502 standard; cDNA: 3311 BP.
 XX
 AC AAI29502;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE C902P determined cDNA sequence.
 XX
 KW Human: immunotherapy; diagnosis: colon cancer; colon tumour;
 KM immunogenic; gene therapy; vaccine; colonic cancer; ss.
 XX
 OS Homo sapiens.
 PN WO200149716-A2.
 PD 12-JUL-2001.
 XX
 PF 29-DEC-2000; 2000WO-US35596.
 XX
 PR 30-DEC-1999; 99US-0476296.
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.
 XX
 PA (CORI-) CORIXA CORP.
 PI Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;
 PI King GE, Wang T, Jiang Y;
 DR WPI: 2001-441847/47.
 XX
 PT Colon tumor associated proteins and nucleic acids useful for the
 PT prevention, diagnosis and treatment of colonic cancer -
 PS
 PS Claim 2; Page 425-426; 472pp; English.
 CC The present invention describes colon tumour associated proteins (I) and
 CC the polynucleotides (II) that encode them. (I) have cytosolic activity.
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate colon tumour associated protein (TCAP)
 CC expression, such as colonic cancer. For example, (I) and (II) may be
 CC used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of TCAPs by expressing inactive proteins or to supplement the
 CC patients own production of them. Additionally, (II) may be used to
 CC produce the TCAP proteins, by inserting the nucleic acids into a host

CC cell culturing the cell to express the protein. (II) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC polymerase chain reaction (PCR) and hybridisation assays to detect and
 CC quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. (I) may
 CC also be used as antigens in the production of antibodies against TCAPs
 CC and in assays to identify modulators of TCAP expression and activity.
 CC Anti-(I) antibodies and antagonists may also be used to down regulate
 CC TCAP expression and activity. The anti-(I) antibodies may also be used
 CC as diagnostic agents for detecting the presence of TCAPs in samples
 CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
 CC and AAM24494 to AAM24523 represent nucleotide and amino acid sequences
 CC given in the exemplification of the present invention.
 XX
 SO Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other:
 Query Match 99.3%; Score 1671; DB 22; Length 3311;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1682; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 OY 1 AACAAAGTGTGCTCATCATCCACAGAGCGCTTGGGCGCTCGACGCTCAAGACTAG 60
 |||||||
 Db 1628 aacaaagtggtgcatacatccacacagtcgttggggcctcgcagctcaagaactag 1687
 OY 31 AGAGCTGTCCAAATGACAGAGAGGTTTACAGACATATGCTTCAGATCAAGTTAGACA 120
 |||||||
 Db 1638 aggagctgtccaaaatgacagaggggtttacagacataatgctcagatcagtagaca 1747
 OY 121 ATGGCTCATGATGCTTTTGGGCGCTTTCATCAGCAAGAAATGACGCTCTCTAGCGCT 180
 |||||||
 Db 1748 atggcctcatgagcttcttgggcccccttcatacagaaatgagagcttctcctaagcct 1807
 OY 131 CCATCCAGCTTGAGAGTAAAGGATTAACCTCCAGAACAGCAGAGATGAATGGACAG 240
 |||||||
 Db 1808 ccatccagcttgagagtaaggaatuaaccctccagaaagcagtgatgataatgacag 1867
 OY 241 TGATCGTGAAGACAGCAGCTGGGAAAGACACTTTGTTTATCAGCTGACCAACGACG 300
 |||||||
 Db 1868 tgatcgtgagacagcagctgggaaagacacttcttctatccctgacaagcagc 1927
 OY 301 CTCGCCAATCTCTTCTCGGAGTCCCATGACAGCAACAGCGCTTTCAGTGGACA 360
 |||||||
 Db 1928 ctccccaatctcttctcggagatccagtgacagaaagcaggttcttctgtagtgcga 1987
 OY 361 AAACACCAAAATGCGCTACCTCCAAATCCAGCATGCTTAAGTTGGACTTGAAT 420
 |||||||
 Db 1928 aaacacaaaatgcttcttctcagctcccaatccagcaggttcttctgtagtgcga 2047
 OY 421 ACAGTTCGACAGCAAGCTACCAACCTTACCTGACGCTCAGCTCCGCTGCTCAATG 480
 |||||||
 Db 2048 acagtcgacagcaagctaccaaacttgacctgacgtcagctccgctgctcagtg 2107
 OY 431 CTACCTGCTCCATTTACAGTACGACTTCCAAACAGCAAGAGCAGCAAGCAATGCCCA 540
 |||||||
 Db 2108 ctacctgctcccaatcagagtgacttccaaaacgaagcagcagcaaatccca 2167
 OY 541 GCCCTGTGTTATTTATGCAATATTTGCGCAAGAGCGCTCCCAATTTCTCAAGGCGACTG 600
 |||||||
 Db 2168 gccctgtgtatttatgcaaatatctgcgaagcgctccccaattctcagggcagtg 2227
 OY 601 TCACAGCCCTGATTAATGATGCAATGGAATAACAGTTACTCTTGGAACTACTGGAATG 660
 |||||||
 Db 2228 tcacagccctgattgaatcagatgaaatgaaacagcttacctggaactactggaatg 2287
 OY 661 GAGCAGTGTGATGCTTCTAAGATGAGCGGTCTACCAAGATTAATTTTACCACTTATG 720
 |||||||
 Db 2288 gagcagtgatgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 2347
 OY 721 ACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 |||||||
 Db 2348 aacagatgatacagtgtaaaaaatgagccttgggaggttaacagcagcagac 2407

QY 781 GGAGAGTATACCCGAGAGTGGAGCACTGACATGACTGCTGGATTGAGAAATGATG 840
 DB 2408 ggaagatgatacccccagagatgagacactgacatactgagctgagatgagaaatg 2467
 QY 841 AATATCAATGGAATCCACAGACCTGAAATTAAGTATGATGATCAACACAAACAG 900
 DB 2468 aatacaaaatggaatccacagacactgaaatataaagatgagatgacacaaag 2527
 QY 901 TGTGTTCAGCAGAACATCTCCGGAGGCTCATTTGTGGCTTGTGATGTCCTCAAGTCT 960
 DB 2528 tgtgttcagcagaatactcggagagctcatttggcttctgattgctccaaatgctc 2587
 QY 961 CCATACCTGATCTCTTCCACACTGGCCAAATACCGACCTGAAGCGGAAATTCACGGG 1020
 DB 2588 ccataactgatctctccacactgacacaaatcacgacactgaaagaggaattcaacg 2647
 QY 1021 GGAGTCTCATTTATCTGATCTGGACAGCTCCCTGGGATGATTTATGACCATGAAACGCTC 1080
 DB 2648 ggaagctcatcaatctgacttggacagctcctgggagatgattatgacatgagacagctc 2707
 QY 1081 ACAGATATATCATTTGAAATAGTACAGATATTTGATGATCAGAGACAAAGTCAATGAAT 1140
 DB 2708 acaagatatactatcgaataagatgacagatcttctgactcgaagacaaatgcaatg 2767
 QY 1141 CTCTTCAATGAAATGATCTGCTCTCATCCCAAGAGCAACTCTGAGGAAGTCTTT 1200
 DB 2768 ccttcaagatgatactactgctctcatcccaagagaaagcaactctgaggaagctctt 2827
 QY 1201 TGTTTAAACCAAAAACATTACTTTGAAAATGGACAGATCTTTTATGCTATTCAGG 1260
 DB 2828 tgtttaaaccagaaacattacttggaaaatgagcagatcttctcattcagctcag 2887
 QY 1261 CGTTGATTAAGTGCATGTAATCAAAATATCAAAATGACAGATCTTTGTTTA 1320
 DB 2888 cgtgtgataagtgatcgtcgaataatccaaatctgacagatcttcttctt 2947
 QY 1321 TTTCTCCACAGACTCCGCCAGAGACACTGATCTCTGATGAAACGCTGCTCTTGTCTTA 1380
 DB 2948 ttctcccaagactccgcagagacacactagctcgtgaaacgctcgtccttgcctta 3007
 QY 1381 AATTCATATCAACAGACACTCTCGGACATTCATTTTAAATATGATGAGACGGA 1440
 DB 3008 aatctcatatcaacagacacactccgacatctacatcttaaaatataatgagatg 3067
 QY 1441 TAGGAACATGAGCTGCTGTAATAGCTAGGGCTGAATTTTGTGAGATTAATTAATAA 1500
 DB 3068 taggagaactgcaagctgctcaatagctagggcgaatttctgcaataataataaa 3127
 QY 1501 TCATTGATCTTTTGTGATTAATAATTTTCTAAATGATTTTGTAGACTTCTGTAGG 1560
 DB 3128 tcatctatcc-ttttttgattataaatttcttaaaatgattcttagacttctgtag 3186
 QY 1561 GGGCGATATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
 DB 3187 gggcgatacaataatgataatgataatgataatgataatgataatgataatgataatg 3246
 QY 1621 AATGATTAATGATTTTGAATCTCTGAGGGGCGATTAATAATGATTAATGATTAATGAT 1680
 DB 3247 ataataaataatgattttagacttctgtagggcgataataataatgctaaacaactg 3306
 QY 1681 GTA 1683
 DB 3307 gta 3309

RESULT 4
 ID AAH34879 standard; cDNA: 2854 BP.
 AC AAH34879;
 XX
 DT 03-SEP-2001 (first entry)
 XX

DE Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 1; ss.
 OS Homo sapiens
 PN NC_00122920-A2.
 PD 05-APR-2001.
 XX 26-SEP-2000; 2000MO-US26524.
 XX 29-SEP-1999; 99US-0157137.
 XX 03-NOV-1999; 99US-0163280.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PI Rugen SM, Barash SC, Birse CE, Rosen CA;
 DR WPI: 2001-235357/24.
 DR P-PSDB: AAG75474.
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PS useful for preventing, diagnosing and/or treating colorectal cancers -
 C1: 1: Page 3462-3463; 9803pp; English.
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;
 Query Match 90.8%; Score 1528.4; DB 22; Length 2854;
 Best Local Similarity 99.68; Pred. No. 0;
 Matches 1532; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 131 AACAAGTGTGTCATCATCACACAGTCTGTTGGGGCCCTTCATCAGGAATGAGCTGCTCAGCGCT 180
 DB 1311 aacaagtggtgctcatcatcacacagctcgcttggggcccttcaggaatgagctgctcagcgct 190
 QY 1311 aacaagtggtgctcatcatcacacagctcgcttggggcccttcaggaatgagctgctcagcgct 190
 DB 1311 aacaagtggtgctcatcatcacacagctcgcttggggcccttcaggaatgagctgctcagcgct 190
 QY 1311 AAGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAATGATTCAGACAA 120
 DB 1371 aggaagctgtccaaatgacagaggtttacagacatattgcttcagatcaagttcagaaca 1430
 QY 1371 AATGGCTCATGATGATGCTTTGGGGCCCTTCATCAGGAATGAGCTGCTCAGCGCT 180
 DB 1411 atggcctcatgattgatttggggcccttcaggaatgagctgctcagcgct 190
 QY 1411 CCATTCAGCTTGAAGATTAAGGATTAACCTCCAGAACAGCCAGTGAATGATGACAG 240
 DB 1411 ccatcagcttgaagattaaagattaaacctccagaaacagcagtgatgaatgacacag 1550
 QY 241 TGAATGAGACAGACCCGTTGGAAAGACAGCTTTGTTCTTATACCTGACCAACGACG 300

Db 1551 tgaatcgtagacagcagc tgg9aagaagacactlgtlctlatlcaactcgtgacaacgcagc 1610
 QY 301 CTCGCCAATCTCTCTGGGATCCAGTGGACAGAGAGCTGGCTTGGAGTGGACA 360
 Db 1611 ctcccaaatccctctcctggatccagtgagacagaagaagtgagcttctgaaagtgagaca 1670
 QY 361 AAAACACCAAAATGGCTCTACTCTCAAAATCCAGGATCTGTAAGTGTGGCACTTGGAAAT 420
 Db 1671 aaaaacacaaaatggcctactcccaatcccaagcagctgtcaagttgtgcaacttgaaat 1730
 QY 421 ACAGTCTGCAAGCAAGCTTCACAAACCTTGACCTGACCTGCTGCTGCTGCTCAATG 480
 Db 1731 aacagctctgcaagcaagcttcaacaaaccttgaccctgactcaagctccgctgctgctcaatg 1790
 QY 481 CTACCCCTGCTCCCAATTCAGTACTCCAAAAGACAGAGACGACGCAAAATTCGCCCA 540
 Db 1791 ctacccctgctcccaatctacagtgacttccaaaagaacagacaccccaatcccca 1850
 QY 541 GCCCTCTGATGTTTATTCGAATATTCGCCAAGAGCTCCCAATTCAGAGGCTGAGT 600
 Db 1851 gccctctgtagtttatgcaaatatctgcgaagagcctcccaatcttcagagcagtg 1910
 QY 601 TCACAGCCCTGATTCGAATTCAGTGAATGGAATAACAGTTACCTTGAGACTACTGATTAATG 660
 Db 1911 tcacagccctctgacttgaatcagtgaaatgaaacagttacacttgaaactcgtgaaatg 1970
 QY 661 GAGCAGGCTGATGCTCTCTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 720
 Db 1971 gagcagagtgctgagtgctcctaagagtgagtgctcctaagagtgagtgctcctaagagtgag 2030
 QY 721 ACAGCAATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 780
 Db 2031 aacagcaatgtagagtagacagtgaaagtgagtgagtgagtgagtgagtgagtgagtgag 2090
 QY 781 GGAGAGTATACCCAGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAG 840
 Db 2091 ggaagagtagatcccaagagagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2150
 QY 841 AAATCAATGCAATGCACCAAGAGCTGCAAAATTAAGAGTGAAGTGAAGTGAAGTGAAGT 900
 Db 2151 aaatcaatgtagatcccaagagagtgagtgagtgagtgagtgagtgagtgagtgagtgag 2210
 QY 901 TGTGTTTACAGCAAGATCTCTGGAGGCTCTATTTGTGGCTTCTGATGCTCCAAATGCTC 960
 Db 2211 tgtgtttacagcaagatccctcctggagtgagtgagtgagtgagtgagtgagtgagtgag 2270
 QY 961 CCATACCTGATCTCTCCCACTGCGCAAAATACCGAGCTGAGAGGCGGAAATTCACGGGG 1020
 Db 2271 ccataccctgactctctcccaacttgcgaatcaccagcctgagagcggaatcaccagcg 2330
 QY 1021 GCAGTCTGATTAATCTGACTTGGAGAGCTCTGGAGATGATTAATGACATGAGACAGCTC 1080
 Db 2331 gcaagtctaatctatctgtcgtgagagctcctcgtgagtgagtgagtgagtgagtgagtgag 2390
 QY 1081 ACAGTATATCTCTGGAATTAAGTACAGTATCTTGAATCTGAGAGCAAGTCAATGAAT 1140
 Db 2391 acaagtatactatctcgaataagtagaagtagatcttgcagtagagacaagtcaatgagat 2450
 QY 1141 CTCCTCAGTGAATCTCTGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1200
 Db 2451 cctctcaagtgatactactcgtctcctcctcctcctcctcctcctcctcctcctcctcct 2510
 QY 1201 TGTTAACCAAGAAATTAATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260
 Db 2511 tgtttaaaccaagaaatctacttcttgaataagtagaagtagatcttctcctcctcctcct 2570
 QY 1261 CTGTTGATTAAGTGAATCTGGAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320
 Db 2571 ctgtgataagtgatctgagtagaagtagaagtagaagtagaagtagaagtagaagtagaagtag 2630
 QY 1321 TTCCGCAAGAGCTGCGGAGAGAGACAGTACGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
 Db 2631 ttccgcacagagctcgc 2690

QY 1381 ATATCATATCAACAGCAGCATCTCTGCGATTCACATTTTAAATATGTAAGTGA 1440
 Db 2631 atattcataatacaagcaccatctctgcatcacaattttaaatttgggaagtga 2750
 QY 1441 TAGGAACTGAGCTGTAATAGCTAGGCTGTAATTTTGTGAGATTAATTAATAA 1500
 Db 2751 taggaactgagctgagtcacatagcctagagctgagatttgcagataataataaa 2810
 QY 1501 TCATCATCTCTTTTGTGATTAATTAATTTTCTTAA 1538
 Db 2811 tcatcatccttcttctgattataaaaaaaaaa 2848
 RESULT 5
 AAF81787
 ID AAF81787 standard; cDNA; 2854 BP.
 AC AAF81787;
 XX
 XX
 DT 12-JUN-2001 (first entry)
 XX
 DE Human secreted protein gene 1 SEQ ID NO:11.
 XX
 KW Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
 KW dermatological; immunosuppressive; anti-inflammatory; anti-HIV;
 KW immunostimulant; cytoskeletal; cardiac; vascular; anti-angiogenic;
 KW epithelium; neuroprotectant; nootropic; anticonvulsant; vaccine;
 KW Alzheimer; antiparkinsonian; antileukemia; antileukemia; gene therapy;
 KW immune disorder; hyperproliferative disorder; cardiovascular disease;
 KW cancer; angiogenic disorder; neurological disorder; infectious disease;
 KW wound healing; regeneration; chemotaxis; chromosome 1; ss.
 XX
 OS Homo sapiens.
 XX
 PN WC200112775-A2.
 XX
 ED 22-FEB-2001.
 FD 16-AUG-2000; 2000WO-US22325.
 PF 17-AUG-1999; 99US-0149182.
 PR
 XX
 FA (HMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ni J, Florence KA, Piscella M, Wei P, Baker KP;
 FI Biese CE, Young PE, Komatsoulis GA, Moore RA, Soppet DR;
 XX
 WP: 2001-14750/15.
 DR P-ESDB: AAB74733.
 XX
 PT Nucleic acids encoding 25 human secreted polypeptides, useful for
 PV preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 IT disease and diabetic retinopathy -
 TX
 XX
 PS Claim 1: Page 441; 485pp; English.
 CC AAF81787 to AAF81817 encode the human secreted proteins given in AAB74733
 CC to: AAB74772. Human secreted proteins can have activities based on the
 CC tissues and cells they are expressed in. Example of activities include:
 CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;
 CC anti-inflammatory; anti-HIV; immunostimulant; cytoskeletal; cardiac;
 CC vascular; anti-angiogenic; epithelium; neuroprotectant; nootropic;
 CC anticonvulsant; Alzheimer; antiparkinsonian; antileukemia; and
 CC vaccine. Human secreted proteins can be used in gene therapy and
 CC vaccine. Human secreted protein nucleotide sequences (NMI) and proteins
 CC associated with inappropriate polypeptide expression. For example, NMI
 CC and PEP1 may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patients genome
 CC that affect the activity of proteins by expressing inactive proteins or
 CC to supplement the patients own production of polypeptides. Disorders that
 CC may be prevented, diagnosed and/or treated include immune disorders,

hyperproliferative disorders (e.g. cancers), cardiovascular diseases, angiogenic disorders, neurological disorders, infectious diseases and/or CC for promoting wound healing, regeneration and/or chemotaxis. AAF8178 to CC AAF8176 and AAF81732 represent sequences used in the exemplification of CC the present invention.

XX Sequence 2854 BP: 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match 90.8%; Score 1528.4; DB 22; Length 2854;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AACAAAGTGGTGCATCATCACACAGTGGCTTTGGGCGCTTGTGACGTCAAGAACTAG 60
DB 1311 acaaaagtgtgcacatcacacacagctcgttgggccccttgagctcaagaactag 1370
QY 61 AGGAGCTGCCAAAATGACAGAGGTTTACACATATGCTTACATCAAGTTCAAGACA 120
DB 1371 aggaagctgtccaaaatgacagagaggtttacagacatattgttcaagttcagaaca 1430
QY 121 ATGGCCCTCATGTATGCTTTGGGCGCTTTCATCAGAAATGAGCTGTCTCAGCGCT 180
DB 1431 atggccctcatgtatgcttttgggccccttcaacagaaatgagctgtctcagcgct 1490
QY 181 CCATCCAGCTTGAAGTAAAGGATTAACCTCCAGAACGCCAGTGAATGAGCAG 240
DB 1491 ccacacagcttgagagtaagagattaacctccagaaagccagtgatgaaatgagcaag 1550
QY 241 TCAATGTGACAGCAGCCGTGGGAAAGACACTTTGTTTCTTATCAGCTGACAGCAGC 300
DB 1551 tcatgtgtgacagcagccgtgggaaagacactttgttcttcaacttgacaagcagcagc 1610
QY 301 CTCGCCAATTCCTTCTGTGGATCCAGTGCAGACAGAAAGTGGCTTTGTAGTGAGCA 360
DB 1611 ctcgccaatctcttctgtggatccagtgctcagagaaagagtgagcttcttgagagca 1670
QY 361 AAAACACAAAATGGCTTACCTCCAAATCCAGGACTTCTAAGTTGGCACTTGGAAAT 420
DB 1671 aaaacacaaaatggcttaccctccaaatccagcacttgagtgacacttgaaat 1730
QY 421 AAGAGTGTGAACAGCAAGCTCAAAACCTTGACCTGTACAGTCCGCTGCTCAAGT 480
DB 1731 aagagtgtgaacagcaagctcaaaaccttgacctgactgtcaagctccgctgcaaat 1790
QY 481 CTACCTGCTCCAAATTTACAGTACTTCCAAAAGCAAGACAGACAGCAAAATTCGCCA 540
DB 1791 ctacctgctccaattacagacttccaaaagcaagacagacacacaaatcccca 1850
QY 541 GCCCTCTGTAGTTTATGCAAAATTTGCCAAGAGGCTCCCAATTTCCAGGCCAGTG 600
DB 1851 gccctctgtagtttatagcaaatatctgcgaagagccctcccaattccagggcagtg 1910
QY 601 TACAGAGCCCTGATGAAATGAGTGAATGAAAAACACTTACTTGGAACTACTGATTAAT 660
DB 1911 tcaagagccctgattgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1970
QY 661 GAGCAGGTCTATGCTTAAAGATGAGGCTGTACTCAAGCTATTTTCACTCAACTTATG 720
DB 1971 gagcaggtctatgcttactaagatgaagtggtctactcaaggtatttcaacttaag 2030
QY 721 ACACGAATGTAATGACATGTAATAAGTGGGCTCTGGAGAGATTAAACGACGACAG 780
DB 2031 acacgaatgtatagatagatagatagatagatagatagatagatagatagatagat 2090
QY 781 GGAGAGTGTATCCAGCAGAGTGGAGCACTGTACATACTGGCTGGATTGAGAAATGAT 840
DB 2091 ggagagtgtatccagcag 2150
QY 841 AAATCAATGAGTACACAGACCTGAATTAATTAAGATGATGTTTCAACACAGCAAG 900
DB 2151 aaatcaatgagatcacacaaagactgaataataataagatgatttcaacacagcaag 2210

QY 901 TGTGTTGACAGACATCTCTGGAGGCTCATTTGTGCTTGTATGTCCAAATGCTC 960
DB 2211 tgtgttgacagacatctctggagggctcatgttggcttgtatgtccaaatgctc 2270
QY 961 CCATACCTGATCTTCTTCCACCTGGCCAAATTCACGACCTGAAAGCGGAAATTCAGGGG 1020
DB 2271 ccatacctgatcttcttccacctggccaaattcacgacctgaaagcggaattcacg 2330
QY 1021 GCAGTCTAATTAATGACTTGGAGACAGCTCTCTGGGATGATTAATGACATGAAACGCTC 1080
DB 2331 gcagtctaattaatgacttggagacagctctctgggattgattaatgacatgaacgctc 2390
QY 1081 ACAAGTATATCATGCAATTAAGTACAGTATTTCTTATCTCAGAGCAAGTTCAATGAAT 1140
DB 2391 acaagtatatcatgcaattaaagtacagatatttcttgcctcagagcaagttcaatgaat 2450
QY 1141 CTCTTCAAGTATGATCTACTGCTCTCATCCCAAGAGACCAACTCTGAGGAAGTCTTTT 1200
DB 2451 ctcttcaagtatactactgctctcatcccaagagaccaactctgaggaagctcttt 2510
QY 1201 TGTTTAAACAGAAACATTAATTTGAAATGSCACAGATCTTTTCAATTCATTCAGG 1260
DB 2511 tgttttaaacagaaacattaatcttgaataatgagcaagatcttccatgcatcag 2570
QY 1261 CTGTTGATTAAGTGTGATCTGAAATCAGAAATATCAACATTCAGCAGATCTTTGTTT 1320
DB 2571 ctgttgatgaagtgatctgaaatcagaatataccaacttgagagatcttctgttta 2630
QY 1321 TTCTCTCAGACACTCCGCCAGAGACACACTAGTCTGATGAACGCTGCTCTTGTCTTA 1380
DB 2631 ttctctcagacctccgccagagacactagctgatagaagctgctctgttctcta 2690
QY 1381 AATTCATATCAACAGACACACTCTGCAATTCATCAATTTTAAATTAATGAGAGTGGA 1440
DB 2691 aatctcatatcaaacagacacctctgcatctacatcttaataatgaggaagtgga 2750
QY 1441 TAGGAGAACTGACAGCTGCAATTAAGCTAGGCTGAATTTTGTACAGATAAATAATTA 1500
DB 2751 taggagaactgacagctgcaattaaagctagggcctgaattttgttcagataataata 2810
QY 1501 TCATTCATCTCTTTTGTGATTTAAATTTTCTAAAT 1538
DB 2811 tcattcatctcttcttattatataaaaaa 2848

RESULT 6
AAH33285
ID AAH33285 standard; cDNA; 2867 BP.
AC AAH33285;
DT 03-SEP-2001 (first entry)
XX Human colon cancer antigen encoding cDNA SEQ ID NO:341.
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; chromosome 1; ss.
XX Homo sapiens;
IN WC200122920-A2.
IN 05-APR-2001.
XX 26-SEP-2000; 2000WO-US26524.
XX 29-SEP-1999; 99US-0157137.
XX 03-NOV-1999; 99US-0163280.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rajen SM, Barash SC, Birse CE, Rosen CA;

DR WPI: 2001-235357/24.
DR P-PSDB: AAG73854.
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX
PS Claim 1: Page 2452-2453; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing P.
CC Inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SO Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other:

Query Match 90.7%; Score 1526.8; DB 22; Length 2867;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 AACAAAGTGGTCCATCATTCACACAGTCGCTTTGGGGCCCTCTCAGCTCAAGACTAG 60
DB 1315 aacaaagtgggtccatcatccacacagtcgcttggggccctctcagctcaagactag 1374
OY 61 AGGAGCTGCGCCAAATGACAGAGGTTTACACATATGCTTCAGATCAAGTTGCAACA 120
DB 1375 aggaagctgctccaaatgacagaggtttacacatatagtcttcacagttcaagttcaaca 1434
OY 121 ATGGCCTATGATGCTTTTGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCGCT 180
DB 1435 atggcctatgattgcttttggggcccttcttcacagaaatgagctgtctctcagcgct 1494
OY 181 CCATCCAGCTTGAAGTAAAGGATTAACCTCCAGAACAGCCAGTGGATGATGGACAG 240
DB 1495 ccatccagcttgaagttaaggttaaacctccagaaacagccagtggatgattggacag 1554
OY 241 TGATGTGGAGACGACCGGGAAGAGACACTTTTCTTTCACGTCGACAGCAGC 300
DB 1555 tgatgtggagacgacccgggaagagacacttttcttcttcacgtcgacagcagc 1614
OY 301 CTCGCCAAATCTTCTTGAGATCCCAAGTGACAGAGCAGAGGCTTTTGTAGTGACA 360
DB 1615 ctccccaatctctctctggagatcccaagtgcagagagcagagcttttgtatgacac 1674
OY 361 AAAAACCACAAATGGCTTACCTCCAAATCCAGCATTTGCTTAAGTTGGCAATTGGAAAT 420
DB 1675 aaaaaccacaaatggcttaccctccaaatccagcatattgagcttggcaaat 1734
OY 421 ACAGCTGCAAGCAGACGACAAACCTTGACCTGACTCTACGTCGCGCTGGTCCAAATG 480
DB 1735 acagctgcaagcagacgacaaaccttgacctgactctacgctcgcgctggctccaaatg 1794
OY 481 CTACCTGCTCCAAATTAAGTACTTCCAAAGAACAGACAGACACCAATTCCTCCCA 540
DB 1795 ctacctgctccaaatgaagtacttccaaagaaacagacagacaccaatctcccca 1854
OY 541 GCCCTCTGTTATTTATGAAATATTCGCCAAGAGCCCTCCCAATTTCTCAGGCGCAGTG 600
|||||

DB 1885 gccctctgtagttatgatgcaaatatctgcgaaggagccctcccaattctcaaggccagtg 1914
OY 601 TCACAGCCCTGATTTGAATCAGTGAATGGAATAAAGAGTACTTCTGAACTACTGATATG 660
DB 1885 tcacagccctgattgaaatcagtgatgaaatgaaataacagttactcttgaaactctgataatg 1914
OY 651 GAGCAGGTGCTGATGCTACTAAGATGAGCGGTGTCTACTCAAGTATTTCACAACTTATG 720
DB 1975 gagcaggtgctgattgaaatcagtgatgaaatgaaataacagttactcttgaaactctgataatg 2034
OY 721 ACAGCAATGCTAGATTAAGCTTAATAAGTCGCGCTCTGCGAGGAGTAAACCAACCAAGAC 780
DB 2035 acacgaatgctagattagatagatgaaatgaaatgaaataacagttactcttgaaactctgataatg 2094
OY 781 GGAGAGTATGATCCCGACAGAGTGGAGAGCTGATACATACCTGCGTGAATGAGATGATG 840
DB 2095 ggagagatgattcccgacagagtgagagagctgatacatctcgcttgatgataatgagtg 2154
OY 841 AATATCAATGGAATCCACCAAGACCTGAAATTAATAGATGATGTTCAACACAGCAAG 900
DB 2155 aatataatggaatccacccaagacctgaaatataatagatgattgatacaacaagaag 2214
OY 921 TGTGTTTACGAGAAACATCTCCGGAGGCTCATTTGTGCTTCTGATGTCCTCAATATGCTC 960
DB 2215 tgtgtttacgagaaacatctccggaggctcatattgtgcttctgattcccaatgctc 2274
OY 961 CCATACCTGATCTCTTCCACCTGCGCAATACACGACCTCAAGCGGAAATTCACGGGG 1020
DB 2275 ccatacctgatctctctccacctgccaatatacacgacctgaaggcggaatctcaaggcg 2334
OY 1021 GCAGTCTCATTAATCTGACTTGGACAGCTCCTGGGATGATTAATGACCATGGAACAGCTC 1080
DB 2335 gcagctctcatattaatctgacttggacagctcctgggattgattatgacatggaacagctc 2394
OY 1081 ACAAGTATATGATCGATTAAGTAACTAATGATTTGATCTCAGAGACAGTTCATGAT 1140
DB 2395 acaagtatatgattcgattaaactaattgatttgaattctcagagacagttcatgattg 1140
OY 1141 CTCTCAAGTAAATTAATCTGCTCTCATCTCCCAAGGAGCAACTCTGAGAGAGTCTTTT 1200
DB 2455 ctctcaagttaatttaattctgctctcatctcccaaggagcaactctgaggaagctctt 2514
OY 1201 TGTTTAAACGAAATTAATCTTTTGAATATGGACAGATCTTTTCTATTCAGG 1260
DB 2515 tgtttaaacgaaatataattcttttgaatattggacagatcttttctatctcag 2574
OY 1251 CTGTGATTAAGTCTGATGTAATGAAATATCCAAATGACAGATTAATCTTTGTTTA 1320
DB 2575 ctgtgattaaagtctgattgtaattgaaatataccaaatgacagatgattcttgtt 2634
OY 1321 TTCTTCACAGACTCCGCGACAGACACCTAGTCTGATGAACGTCCTCTCTGCTCTA 1380
DB 2635 ttcttcacagactccgcgacagacacctagtcctgataaagctctgctcttctccta 2694
OY 1381 ATATTCAATCAACAGCAGCATTTCTGCTATTCATTTTAAATTTTGGGAAGTGA 1440
DB 2695 atattcaatcaacagcagcatcttctgctatctcatctttaaatttgggaagtga 2754
OY 1441 TAGGAGAACTGAGCTGCAATAGCTGAGGCTGAATTTTGTGATGAATTAATAAT 1500
DB 2755 taggaagaaactgagctgcaatagctgagggctgaatttgtgcaataataataataa 2814
OY 1501 TCATTCAATCTTTTGTGATTAATAATTTCTTAA 1538
DB 2815 tcattcaatcttttgtgattataataataataataa 2852
|||||

RESULT: 7
AAH46121
ID AA46124 standard; cDNA; 2825 BP.
XX
XX AA46124;
XX

DT 11-SEP-2001 (first entry)
 XX Human CLCA1 cDNA, SEQ ID NO:26.
 DE
 XX Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
 KW expression inhibition; antisense therapy; gene therapy;
 KW chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
 SS.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 23..2767
 FT /tag= a
 FT /product= "Human CLCA1"
 FT /trans_except= (pos:476..478, aa:lys)
 FT
 PN WO200138530-A1.
 PD 31-MAY-2001.
 PE 22-NOV-2000; 2000MO-JP08232.
 PR 24-NOV-1999; 99JP-0333479.
 PR 27-APR-2000; 2000JP-0127589.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Nakanishi A, Morita S;
 DR WPI: 2001-355935/37.
 DR P-PSDB: AAB73716.
 XX
 PT New antisense nucleotide, useful for treatment and prevention of
 PT bronchial asthma and chronic obstructive pulmonary disease -
 XX
 PS Example 5: Page 92-94; 104pp; Japanese.
 XX
 CC The invention relates to an antisense nucleotide targeted to the mouse
 CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
 CC relates to an antibody specific for the Gob-5 protein, medical and
 CC diagnostic compositions containing the antisense nucleotide or the
 CC antibody, and methods and kits for screening for compounds which inhibit
 CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 CC The antisense oligonucleotides and antibody are therefore useful for the
 CC treatment and prevention of bronchial asthma and chronic obstructive
 CC pulmonary disease. The present sequence represents human CLCA1 cDNA.
 CC
 XX
 SQ Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;

Query Match 90.2%; Score 1517.4; DB 22; Length 2825;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1521; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 0Y 1 AACAAAGTGGTCATCATCAGACAGAGCGCTTTGGGCCCTGCGAGCTCAAGACTAG 60
 DB 1299 acaaaagtggtgcatcacatcacacagctgcttggggccctcgagcccaagaactag 1358
 0Y 61 AGGAGCTGTCAAAATGACAGAGAGTTTACAGACATATGCTTCAGATCAAGTTCACANACA 120
 DB 1359 aggagctgtccaaatgaaagaggtttacagacatagtcttcaagttcagaaca 1418
 0Y 121 ATGGCTCATTTGATGCTTTTGGGGCCCTTTTCATCAGAGAAATGAGCTGTCTCAGCGCT 180
 DB 1419 atggctcatatgctgttcttggggcccttccatcagaagaatgagctgtctcagaagc 1478
 0Y 181 CCATCCAGCTTGAAGTAAAGGATTAACCTCCAGAAACGCCGTGATGATGAATGGACAG 240
 DB 1479 ccatccagcttgaagtaagattaaacctccagaagccagtgatgaatggacag 1538
 0Y 241 TGAATGTGACAGACCGGTGGGAAAGACATTTGTTTATCACCCTGGACCAACGACG 300
 1519 tgaatgtgacagaccggtgggaaagacattgttcttaccactgagcaacgacg 1598
 301 CTCCCAATTCCTTCCTCTGGGATCCAGTGGACAGAGAGAGTGGCTTTGTGAGACA 360
 DB 1599 ctcccaaatccctccctcctggagccagtgagaaagaagaagtgcttcttgaatgagaa 1658
 361 AAAACACAAAATGAGCTCCTCCAAATCCAGGCAATTCCTAAGTGTGGCACTTGAAT 420
 DB 1659 aaacacaaaatgagctcctccctcccaatccagcttgaagttgagacttggaaat 1718
 431 ACAGTGTGACAGCAGCTCAAAACCTTGAACCTGTACGTCTACGTCCGCTGCTCAATG 480
 DB 1719 acagtgtgaagaagctcaacaacctgacctgtgtcagctccgctgctgctcaatg 1778
 481 CTACCTCGCTCCCAATTTACAGTACCTTCCAAAAGCAACAGACACACCAATTCGCCA 540
 DB 1779 ctacctgctcccaatlaaaglaacttccaaaacagaacagacacagcaaatcccca 1838
 541 GCCCTGTGATGTTTATGCAAAATATTCGCCAAGAGACCTCCCAATTCAGGGCCAGTG 600
 DB 1839 gccctgtgatttattatgcaaatatccgcaagagagctcccaatctcagggccaagt 1898
 601 TCACAGCCCTGATTTGAATCAGTGAATGAAAAACAGTTACCTTGAACCTAGTGAATATG 660
 DB 1899 tcacagccctgattgatacagtgaaatgaaacagtlacttggaaactactgataatg 1958
 661 GACAGAGTGTGATGCTTACTAAGATGACGAGGTCTTACTACAGATTTTACAACTTATG 720
 DB 1959 gacagagtgtgactgtactaaagatgacggtgtctacttcaaatgatttcaacaactatg 2018
 721 ACAGAAATGGTATGATACAGTGAATGAGGAGGCTGTGGAGAGATTAACGACGACGAC 780
 DB 2019 acagaaatggtatgatacagtgataaagtgaggtgtgtggagagatgaacgacgac 2078
 781 GGAGAGTATACCCAGCAGAGAGTGGAGACTGTACATPACCTGGCTGATGAGATGATG 840
 DB 2079 ggagagtatacccccagagagtgagagcactgtactactgtgtgataatgagatg 2138
 841 AATTAACATGCAATCCACAGACAGACGCTGAATTAAGTGAATGATGATGATGATGATG 900
 DB 2139 aatlaacatggaatccacacagacccggaatlaaagatgattgttcaacaacaag 2198
 901 TGTGTTTACAGCAATATCCTCGGAGGCTCATTTGTGCTTGTGATGATGATGATGATG 960
 DB 2199 tgtgtttacagcaaatatcctcggaggtcatttggcttctgtgtgtccaaatgctc 2258
 961 CCATPACCTGATCTCTTCCACCTGGCCAAATCACCAGCTGAAGCGGAAATTCACGGGG 1020
 DB 2259 catacctgtctcttccaccctggccaatcacgacatcgaagcggaatccaacg 2318
 1021 GCAGTCTCATTAATTCGACTGTGACAGCTCCTCGGGATGATTTATGACATGGAACGCTC 1080
 DB 2319 gcagtctcataatctgacttggacagctctcgggagatgattatgaccatgagacagctc 2378
 1081 ACAAGTATATCATTCGAATTAAGTACAAATTTTATATCTCAGAGCAAGTGAATGAAT 1140
 DB 2379 acaagtataatcatcgaataagtaacagatcttctatctcagagagaagttcaatgaat 2438
 1141 CTCTTCAAGTAATACTACTGCTCATCCCAAAAGAGCAACCTGAGGAAGTCTTTT 1200
 DB 2439 ctcttcaagtaactactgctctcatcccaagaagcaactctgaggaagctctt 2498
 1201 TGTTTAAACAGAAACATTTACTTTTGAATAATGGACAGATCTTTTCAATTCATTCAGG 1260
 DB 2499 tgtttaaacagaaacattacttcttgaataatggcagagctcttcttcatcttcaatcag 2558
 1261 CTGTTAATAGTGTGATGATGAATCAAAATATCAACATTTGACAGAGTATCTTTGTTTA 1320
 DB 2559 ctgttataagtgatctgaaatcagaataatccaaatctgacagagatctcttctt 2618
 1321 TTCTCCACAGACTCCGCCAGAGACACTAGTCTGATGAAGAGTCTGCTCTTGTCTTA 1380

Db 2619 ttccctcacagactccgcagagacactagctcgtatgaacgctcgtcccttgccta 2678
 QY 1381 ATATTCAATTAACAGACGACCTTCTGGCATTCATTAAATAATATGGAAGTGA 1440
 Db 2679 atattcatacaagaacacccatccctgcacatttaaaatlaatggaagtgga 2738
 QY 1441 TAGGAGAACTGACGCTGATACAGCTAGGCTGATATTTGTCAGATTAATAAATGAA 1500
 Db 2739 tagagagactcagctgcacatagcctagggctgaattttgcaagataataataa 2798
 QY 1501 TCATTACCTCTTTTGTGATTATAA 1527
 Db 2799 tcattcattccttttttgattataaa 2825

RESULT 8
 ID AAF81927 standard; cDNA; 2745 BP.
 XX AAF81927;
 DT 13-JUN-2001 (first entry)
 DE Human ICACC-1 nucleotide sequence.
 XX
 KW ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
 KW Interleukin 9 induced calcium activated chloride channel; IL-9;
 KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
 KW anti-inflammatory; immunomodulatory; cystic fibrosis;
 KW inflammatory bowel disease; autoimmune disease; ss.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..2745
 FT /*tag= a
 FT /product= "ICACC-1"
 FT /note= "IL-9 induced calcium activated chloride channel"
 XX
 PN MO9944620-A1.
 PD 10-SEP-1999.
 PF 03-MAR-1999; 99WO-US04703.
 PR 03-MAR-1998; 98US-0076815.
 PA (MAGA-) MAGANIN PHARM INC.
 PI Holroyd KJ, Levitt RC, Maloy WL, Louahed J, McLane M;
 PI Nicolaidis NC, Zhou Y, Dong Q;
 DR MPI: 1999-550979/46.
 DR P-PSDB; AAB74824.
 PT New nucleic acid encoding calcium activated chloride channel, used to
 PT identify, e.g. specific modulators for treating atopic allergy -
 PS Claim 1: Fig 4B: 75bp; English.

The present sequence encodes the human interleukin 9 (IL-9) induced
 calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
 have anti-allergic, anti-asthmatic, anti-inflammatory and
 immunomodulatory activities. Compounds (A) that downregulate ICACC are
 used to alleviate asthma (or more generally atopic allergy), while those
 (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
 inflammatory bowel disease (IBD) (or other autoimmune diseases).
 CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma
 CC (increased levels) or IBD (reduced levels), also for monitoring
 CC treatment of these conditions. The ICACC proteins can be used:
 CC (1) to raise specific antibodies (Ab), useful: (a) as immunosay
 CC reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
 CC (A) to disrupt bonding between ICACC and its ligands; and (iii) to

CC identify modulators and binding partners. ICACC polynucleotides can be
 CC used to generate transgenic animals or recombinant cells, used to screen
 CC for antagonists, also as a source of therapeutic antisense agents or
 CC diagnostic probes (for quantifying mRNA expression, e.g. for
 CC identification of modulators).
 XX
 SO Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;
 Query Match 87.3%; Score 1469; DB 20; Length 2745;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1381 AACAAAGTGTGCTCCATCATCCACAGCTGCTTTGGGCGCTTCAGCTCAAGACTAG 60
 Db 1277 acaaaagtggtcattatccacacagtcgtttgggacctcgcagctcaagaactag 1336
 QY 61 AGAAGCTGTCCAAATGACAGAGGCTTTACAGACATATGCTTCAGATCAAGTCAACA 120
 Db 1387 aggaagcttccaaaatgacaggaagttacagacatattctcagatcaagttcagaca 1396
 QY 121 ATGGCTTATGATGCTTTTGGGCGCTTTCATCAGAAATGAGCTGTCTCAGGCT 180
 Db 1387 atggcttattgactgttttgggaccttccatcaagaatgagctgtctcagagct 1456
 QY 131 CCATCCACCTTTGAGAGTAAAGGATTAACCTTCAGAACAGCAGTGTGATGAATGGACAG 240
 Db 1497 ccattcagcttgagagtaaggatttaaccctcagaacagcagtgatgagtcagag 1516
 QY 211 TGATGCTGAGACAGACCGTGGGAAAGACACTTTGTTTATACCTGACACAGCAGC 300
 Db 1517 tgatgctgagacagcagctgggaaagacacttgtttcttataccctggacaacgagc 1576
 QY 301 CTCCCAATCTCTCTCGGAGATCCAGTGAGACAGAAAGTGTGCTTTAGTGACA 360
 Db 1577 ctcccaatctctctcgtggtatcccgatgacagaaagtggtcttgaatgaca 1636
 QY 351 AAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAAGTTGGCACTTGAAT 420
 Db 1637 aaaacaccaaatgcttactcccaatccagcagatgtcaagttgacattggaat 1696
 QY 421 ACAGTCTCAGCAAGCTGCAAAACCTTGACCTGACCTGACCTGCGGCTGCAATG 480
 Db 1637 acagtcgcaagcaagcacaacacttgaccctgacgtgcaagctccggtgctccatg 1736
 QY 481 CTACCTGCTCTCAATTAATGACTGACTTCCAAAGCAAGCAAGCAAGCAAGCAATGCCCA 540
 Db 1757 ctacctgctctcaattacagtgacttccaaaagcaagcaagcaagcaagcaatcccca 1816
 QY 541 GCCCTGTGAGTATTTATGCAAAATATTCGCCAAGAGCTCCCAATTTCTAGGCGCAGTG 600
 Db 1847 gccctgtgattatgcaaatatcgcgaagagcctcccaattctcaggcgagtg 1876
 QY 601 TCACAGCCCTGATGAATGATGATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 660
 Db 1877 tcacagccctgattgaatgattgaaagcaagcaagcaagcaagcaagcaagcaagcaag 1936
 QY 661 GAGCAGGTGCTGATGCTACTAAGATGAGCTGTCTACTCAAGTATTTCACAACTTATG 720
 Db 1937 gagcaggtgctgattgactaagatgagctgtctactcaagtatattcaactattg 1996
 QY 721 ACACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 Db 1937 acacgaatgattgattgattgattgattgattgattgattgattgattgattgattgattg 2056
 QY 721 GGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 Db 2057 ggagagtgattgattgattgattgattgattgattgattgattgattgattgattgattgattg 2116
 QY 841 AATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 Db 2117 aatataatgattgattgattgattgattgattgattgattgattgattgattgattgattgattg 2176

XX Sequence 2931 BP: 867 A; 719 C; 684 G; 661 T; 0 other;

Query Match 55.8%; Score 939.4; DB 20; Length 2931;
Best Local Similarity 76.7%; Pred. No. 1.5e-251;
Matches 1210; Conservative 0; Mismatches 341; Indels 26; Gaps 4;

1 AACCAAGTGGGCGCCATCCACAGACGCTTGGGCGCCCTGACCTCAAGAACTAG 60
1287 agcagagggggccatccatccatacagcggccggcgccgctgagcgaagcctg 1346
61 AGAGAGCTGCCAAATGACAGAGGCTTACAGACATGCTTCAGATCAAGTTGACAA 120
1347 agcagctgccaaatgacagagggcgccgacactcttcgatacagctgaggaaca 1406
121 AMGGCCATATGATGCTTTTGGGCGCTTTCATCGAAGAAAGAGCTGCTGCTGCGCT 180
1407 atggtctgt 1466
181 CCATCCAGCTTGGAGAGTAAAGGATTAACCTCCAGACAGCAGCAGTGAATGACAG 240
1467 ccattccagctgtgagagcggggagtaattcccgaaataacaaatgagatgtcag 1526
241 TGATCGTGGACAGCAGCGTGGGAAAGACACTTTGTTCTTATCACTGACACGACG 300
1527 tgatcgtgagacagctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1586
301 CTCGCCAATCTCTCTGCGGATCCCACTGACAGACAAAGCTGCTTGTAGTGGACA 360
1587 ctccacaaatctctctcggggtccagcggagtgagaaacaaatgttlaaatacaga 1646
361 AAAACACCAAAATGCTACCTACCTCAATCCAGCATGCTGAGTGGACATGGAAT 420
1647 caaccacaaatgt 1706
421 ACAGTCTGACAGCAAGCTCAACAACTTGAACCTGCTGCTGCTGCTGCTGCTGCT 480
1707 acagatctcaagcagctcaagcagctcaagcagctcaagcagctcaagcagctca 1766
481 CTACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
1767 ctacacgt 1826
541 GCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
1827 gccctgttaacagctgt 1886
601 TCACAGCCCTGATTAATCAATGATGAAAGAAAGCTTACCTGGAATGATGATG 660
1887 tcacagcctgt 1946
661 GAGCAGTGGTGGATGCTTAAAGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
1947 gaggcagtggt 2006
721 ACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
2007 atgcaaatgt 2066
781 GGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
2067 agagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2126
841 AAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
2127 aagtaagaatgaaccac 2180
901 TGTGTTTACAGCAATCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
2181 tgt 2240
958 CTCCTATCTGATCTCTTCCACCTGGCCAAATCAACGACGACTGAAGGGGAAATTCACG 1017

2231 ctccatctccagcctcttccaccctgtcaatcaatcagctgaagcagcatccag 2300
1038 GGGGAGCTCATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
2331 ggcagagcctgt 2360
1078 CTCACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1137
2361 ctccaactacatcctcagatgtgacagcagcagcagcagcagcagcagcagcagc 2420
1138 AATCTCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197
2421 cctcactcaagtgagac 2480
1198 TTTTGTAAACAGCAAAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1257
2481 ttgagttgagcctgt 2540
1238 AGGCTGTGATAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1317
2541 agcgt 2600
1338 TTATTCCTCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1377
2631 tcatcccgctcag-----gagccgcccattccggaagacacacacacacacacac 2651
1338 CTATATTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1437
2632 ctgacatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2711
1438 GATAGGAGAACCTGAGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1497
2712 ggcag 2771
1438 AATATTCATGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1557
2772 a-----gtatctcttcaatcagagatcttctaaatgtactttagacttctgt 2823
1538 AGGCGGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1574
2824 acggggcggtatagtaa 2840

RESULT 12
ID AAH46101 standard; DNA; 2739 BP.
XX AAH46101:
XX 11-SEP-2001 (first entry)
XX Mouse Gob-5 coding sequence, SEQ ID NO:3.
XX 11-SEP-2001 (first entry)
XX Mouse Gob-5 coding sequence, SEQ ID NO:3.
XX Mouse Gob-5; murine; goblet cell; human C1CA1 orthologue; drug screening;
XX expression inhibition; antisense therapy; gene therapy; bronchial asthma;
XX chronic obstructive pulmonary disease; antiasthmatic; ds.
XX Mus sp.
XX key 1..2739 Location/Qualifiers
XX CD5 /*tag= a
XX FT /product= "Mouse Gob-5"
XX FT /note= "No stop codon given in the specification"
XX PN WO:00138530-A1.
XX 31-MAY-2001.
XX 22-NOV-2000; 2000WO-JP08232.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:11:15 ; Search time 22700.8 seconds
(without alignments)
1223.073 Million cell updates/sec

Title: US-09-049-696-19
Perfect score: 1683
Sequence: 1 AACCAAGTGTGCATCATC.....AAATCTAAACAACACTGGGTA 1683

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 1472140 segs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank1:
1: gb-ba:*
2: gb-ba:*
3: gb-ba:*
4: gb-ba:*
5: gb-ba:*
6: gb-ba:*
7: gb-ba:*
8: gb-ba:*
9: gb-ba:*
10: gb-ba:*
11: gb-ba:*
12: gb-ba:*
13: gb-ba:*
14: gb-ba:*
15: gb-ba:*
16: gb-ba:*
17: gb-ba:*
18: gb-ba:*
19: gb-ba:*
20: gb-ba:*
21: gb-ba:*
22: gb-ba:*
23: gb-ba:*
24: gb-ba:*
25: gb-ba:*
26: gb-ba:*
27: gb-ba:*
28: gb-ba:*
29: gb-ba:*
30: gb-ba:*
31: gb-ba:*
32: gb-ba:*
33: gb-ba:*
34: gb-ba:*
35: gb-ba:*
36: gb-ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1671	99.3	3311	6	AX193489	AX193489 Sequence
2	1671	99.3	3311	9	AF039400	AF039400 Homo sapi
3	1528.4	90.8	2826	9	AF127036	AF127036 Homo sapi
4	1510.4	89.7	2022	9	AK024970	AK024970 Homo sapi
5	1023	60.8	3079	4	AF095584	AF095584 Sus scrofa
6	942.6	56.0	2937	10	AB017156	AB017156 Mus muscu
7	790.8	47.0	878	6	195746	195746 Sequence 8
8	673.8	40.0	3204	9	AF127035	AF127035 Homo sapi
9	672.2	39.9	1895	9	AK000138	AK000138 Homo sapi
10	667.6	39.7	3221	9	AK000072	AK000072 Homo sapi
11	667.6	39.7	3265	6	AX092338	AX092338 Sequence
12	608	36.1	113764	9	HS1651E10	AL122002 Human DNA
13	608	36.1	164891	2	AL358950	AL358950 Homo sapi
14	594.4	35.3	35278	9	AF039401	AF039401 Homo sapi
15	437.8	26.0	2884	4	BT036445	BT036445 Bos tauru
16	414.4	24.6	3317	4	AF001261	AF001261 Bos tauru
17	412.4	24.5	3137	10	AF115852	AF115852 Mus muscu
18	412.4	24.5	3137	10	BC008147	BC008147 Mus muscu
19	410.8	24.4	3058	10	AF108501	AF108501 Mus muscu
20	408.4	24.3	3022	10	AF047838	AF047838 Mus muscu
21	408.4	24.3	3471	10	AF052746	AF052746 Mus muscu
22	366.6	21.8	3415	9	AF043976	AF043976 Homo sapi
23	320.2	19.0	2820	4	AF001263	AF001263 Bos tauru
24	319.2	19.0	3288	4	AF001262	AF001262 Bos tauru
25	301.6	17.9	2832	9	AF127080	AF127080 Homo sapi
26	301.6	17.9	3604	9	AB026833	AB026833 Homo sapi
27	301.6	17.9	4077	6	AX054697	AX054697 Sequence
28	300	17.8	2970	9	AF043977	AF043977 Homo sapi
29	163	9.7	140718	2	AL356270	AL356270 Homo sapi
30	108.4	6.4	140718	2	AL356270	AL356270 Homo sapi
31	108.4	6.4	124096	2	AL354988	AL354988 Homo sapi
32	90.4	5.4	124096	2	AL354988	AL354988 Homo sapi
33	88	5.2	549	6	AX054715	AX054715 Sequence
34	49.4	2.9	1141	6	AX083744	AX083744 Sequence
35	47.8	2.8	1110	10	AF077303	AF077303 Rattus no
36	46.4	2.8	48110	3	AC087794	AC087794 Caenorhab
37	46.4	2.8	159453	2	AC006732	AC006732 Caenorhab
38	46.2	2.7	180569	2	AL354949	AL354949 Homo sapi
39	46.2	2.7	191147	2	AL354949	AL354949 Homo sapi
40	46.2	2.7	208109	2	AC068047	AC068047 Homo sapi
41	45.6	2.7	99083	3	CEY37A1B	AL023835 Caenorhab
42	45.6	2.7	171317	9	AC020941	AC020941 Homo sapi
43	45.4	2.7	120819	33	AL139225	AL139225 Human DNA
44	45.4	2.7	122543	2	AC093119	AC093119 Homo sapi
45	45.4	2.7	137404	2	AL162402	AL162402 Homo sapi

ALIGNMENTS

RESULT 1
AX193489 3311 bp DNA PAT 15-AUG-2001
DEFINITION Sequence 1056 from Patent WO0149716.
ACCESSION AX193489
VERSION AX193489.1 GI:15211440

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCES

Xu, J., Lodes, M. J., Secrist, H., Benson, D. R., Meagher, M. J.,
Stolk, J. A., King, G. E., Wang, T. and Jiang, Y.
Compounds for immunotherapy and diagnosis of colon cancer and
Patent: WO 0149716-A 1056 12-JUL-2001.

JOURNAL

CORTIXA CORPORATION (US)
FEATURES
source 1. 3311
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 99.38; Score 1671; DB 6; Length 3311;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1682; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 AACAAAGGCGCATCATCCACAGCGCTTGGGGCCCTCGACGCTCAAGACTAG 60
1628 AACAAAGGCGCATCATCCACAGCGCTTGGGGCCCTCGACGCTCAAGACTAG 1687
61 AGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTCAAGACA 120
1688 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTCAAGACA 1747
121 ATGGCTCATGATGCTTTTGGGCGCTTTCATCAGAAATGAGAGCTCTCTCAGCGCT 180
1748 ATGGCTCATGATGCTTTTGGGCGCTTTCATCAGAAATGAGAGCTCTCTCAGCGCT 1807
181 CCATCCAGCTTGAGAGTAAAGGATTAACCTCCAGAACAGCAGTGAATGATGACACAG 240
1808 CCATCCAGCTTGAGAGTAAAGGATTAACCTCCAGAACAGCAGTGAATGATGACACAG 1867
241 TGATGTGACAGACCGCTGGGAAAGACACTTTTCTTATCAGCTGACACAGCGACG 300
1868 TGATGTGACAGACCGCTGGGAAAGACACTTTTCTTATCAGCTGACACAGCGACG 1927
301 CTCCCAATCCTTCTGCGGATCCAGTGGACAGACAGAGTGGCTTTAGTGGACA 360
1928 CTCCCAATCCTTCTGCGGATCCAGTGGACAGACAGAGTGGCTTTAGTGGACA 1987
361 AAAACACCAAAATGCGCTACCTCCAAATCCAGCAGTGTGAGTGGACCTTGGAAAT 420
1988 AAAACACCAAAATGCGCTACCTCCAAATCCAGCAGTGTGAGTGGACCTTGGAAAT 2047
421 ACAGCTGACAGACCGCTGGGAAAGACACTTTTCTTATCAGCTGACACAGCGACG 480
2048 ACAGCTGACAGACCGCTGGGAAAGACACTTTTCTTATCAGCTGACACAGCGACG 2107
481 CTACCTGCTCCAAATGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 540
2108 CTACCTGCTCCAAATGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 2167
541 GCGCTGTGATGATTAATGCAATATTCGCCAAGAGCTCCCAATTCAGAGGCGACG 600
2168 GCGCTGTGATGATTAATGCAATATTCGCCAAGAGCTCCCAATTCAGAGGCGACG 2227
601 TCACAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
2228 TCACAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2287
661 GAGCAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
2288 GAGCAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2347
721 ACACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
2348 ACACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2407
781 GAGAGTGTATACCCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 840
2408 GAGAGTGTATACCCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2467
841 AATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
2468 AATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2527
901 TGGTGTGACAGAAATCCTGCGGAGGCTCATTTGCTTGTGATGATGATGATGATGATG 960
2528 TGGTGTGACAGAAATCCTGCGGAGGCTCATTTGCTTGTGATGATGATGATGATGATG 2587

961 CCATACCTGATCTCTTCCACACCTGGCCAAATACGACCGGAAGGGGAAATTACAGGGG 1020
2588 CCATACCTGATCTCTTCCACACCTGGCCAAATACGACCGGAAGGGGAAATTACAGGGG 2647
1021 GCAGTCTGATTAATCTGACCTTGGACAGCTCTGCGGATGATTAATGACCATGAAACGCTC 1080
2648 GCAGTCTGATTAATCTGACCTTGGACAGCTCTGCGGATGATTAATGACCATGAAACGCTC 2707
1081 ACAAGTATATCATTCGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
2708 ACAAGTATATCATTCGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 2767
1141 CTCTTCAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
2768 CTCTTCAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2827
1201 TGTAAACAGAAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
2828 TGTAAACAGAAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2887
1261 CTCTTCAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
2888 CTCTTCAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2947
1321 TTTCTCCAGAGCTCTGCGGACAGACCTGATGATGATGATGATGATGATGATGATGATG 1380
2948 TTTCTCCAGAGCTCTGCGGACAGACCTGATGATGATGATGATGATGATGATGATGATG 3007
1391 ATATTCATATCAACAGACCATTCCTGCGATTCATTCATTTAAATTAATGATGATGATG 1440
3078 ATATTCATATCAACAGACCATTCCTGCGATTCATTCATTTAAATTAATGATGATGATG 3067
1441 TAGAGAACTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
3068 TAGAGAACTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3127
3128 TCATTCATCC-TTTTGTGATTAATTAATTTCTTAATGATGATGATGATGATGATGATGATG 3186
1561 GGGGATATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
3187 GGGGATATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3246
1621 ATACTAATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
3247 ATACTAATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3306
1681 GTA 1683
3307 GTA 3309

RESULT : 2
AF039400
LOCUS AF039400 3311 bp mRNA PRI 14-DEC-1998
DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA,
complete cds.
ACCESSION AF039400
VERSION AF039400.1 GI:4009457
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 3311)
Gruber A.D., Elble R.C., Ji H.L., Schreier K.D., Fuller C.M. and
Pauli B.U.
Genomic cloning, molecular characterization, and functional
analysis of human CLCA1, the first human member of the family of
Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)

MEDLINE 99047526
2 (bases 1 to 3311)
REFERENCE Gruber/A.D., Elble.R. and Pauli,B.U.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
Cornell University, Ithaca, NY 14853, USA

FEATURES
source
1..3311
Location/Qualifiers

gene
CDS
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/tissue_type="small intestine"
1..3311
/gene="hCLCA1"
352..3096
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPKSSVFILILHLEBALSNLIQLNNNGEIVADENP
EDETLIQIKDMVTOASLYLEATGRKREYFKWVAILPEPTWTKADYRPKLETYKNA
DLVAESTPPNDDEPYLSENGRQAVRCAGILGTNVKRCOGSCYTRKCTENKVTGL
YKGEFEVLOSROTEKASIMFAOHVDSIYEETEGNKNEMAKNOKNOMCNSTBEVI
RDSDEFKTTMTTOPRNPFTSLIOIGRICTVLDNKNESAKRNLNINLNGOGLFL
LQYVLEGSWCVKMTFDSAHVQSELIQINSISDRPTLAKRLPAASGGSTISGLRSA
FTVLRKTPDSEIVLITDSDNTISCFNEVKOSGAIIRTVALGPSAOLBELLSK
MIGGLQYASDOVONNGLIDAFALSSNGAASOSIOLSESLGLONSOMNGVIV
DSTGKIDFLITTTTOPPOILLMDPSGOKOGFVADNKTMAVLIQIGIAVGMKY
SLOASSOTLTLVTSRASNATLPITVTSKTKRDSKPPSLVYANIROGASPLIRA
SVTLIESVNGKTVLELLDNGAGADAKDQVYRFTYDTNNGRSYKVALGSVN
AARVRVPOOSGALYIPGMIDEIOMNPREIKDDVOHONCEPSRSSGSYAS
DVNAPRIPDPPPOITDKKAIHGGSLINLTWTPRGGDTHGTHKTIIRISTILD
LQKFNESLQVNTALIPKANESEVFLPKENITFENGDTFLAIQAVDKVLDSEI
SNIAVSLFLIPQTPPETPSPDETSAPCPNINHSTIPGILHLIMMWIGELQLSIA

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 99.3%; Score 1671; DB 9; Length 3311;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1682; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 AACAAAGTGGTCCATATCCACACAGTGGCTTTGGGCCCCCTGCACTCAAGAACTAG 60
1628 AACAAAGTGGTCCATATCCACACAGTGGCTTTGGGCCCCCTGCACTCAAGAACTAG 1687
61 AGGAGTGCACAAATGACAGAGGTTTACAGACATGCTTCAGATCAAGTCAAGAA 120
1688 AGGAGTGCACAAATGACAGAGGTTTACAGACATGCTTCAGATCAAGTCAAGAA 1747
121 ATGGCTCATTTGATGCTTTTGGGCCCCCTTTCATCAGAAATGAGAGTGTCTCTCAGCGCT 180
1748 ATGGCTCATTTGATGCTTTTGGGCCCCCTTTCATCAGAAATGAGAGTGTCTCTCAGCGCT 1807
181 CCATCCAGTTAGAGTAAGGATTAACCTCCAGAAACAGCAGTGGATGAATGGACAG 240
1808 CCATCCAGTTAGAGTAAGGATTAACCTCCAGAAACAGCAGTGGATGAATGGACAG 1867
241 TGATCGTGACAGACCGGAGGAAAGACACTTTGTTTATACCTGAGCAACAGCAG 300
1868 TGATCGTGACAGACCGGAGGAAAGACACTTTGTTTATACCTGAGCAACAGCAG 1927
301 CTCCTCCAAATCTTCTTCTGGGATCCCAATGACAGAGGAGGAGGCTTTGATGAGACA 360
1928 CTCCTCCAAATCTTCTTCTGGGATCCCAATGACAGAGGAGGAGGCTTTGATGAGACA 1987
361 AAAACACCAAAATGGCTCTCTCCAAATCCAGGCAATTCCTAAGTTGGCAGCTTGAAT 420

1988 AAAACACCAAAATGGCTCTCTCCAAATCCAGGCAATTCCTAAGTTGGCAGCTTGAAT 2047
421 ACAGTGCAGAGCAAGCTCAACAACCTGACCTGACGTGACGTCCCGTCCCAATG 480
2048 ACAGTGCAGAGCAAGCTCAACAACCTGACCTGACGTGACGTCCCGTCCCAATG 2107
481 CTACCTGCTCCCAATTAACAGTACCTTCCAAACGAAAGACACACAGCAATTCCTCCCA 540
2108 CTACCTGCTCCCAATTAACAGTACCTTCCAAACGAAAGACACACAGCAATTCCTCCCA 2167
541 GCCCTCTGATGATTTATGCAAAATATTCGCCAAGAGCTCCCAATTCAGAGGCCAGTG 600
2168 GCCCTCTGATGATTTATGCAAAATATTCGCCAAGAGCTCCCAATTCAGAGGCCAGTG 2227
601 TCACAGCCCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
2228 TCACAGCCCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2287
661 GACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
2288 GACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2347
721 ACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
2348 ACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2407
731 GAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
2408 GAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2467
841 AAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
2468 AAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2527
901 TGGTTTACAGCAAAATATCTGCGGAGGCTGATTTGGCTTGTGATGATGATGATGATGATG 960
2528 TGGTTTACAGCAAAATATCTGCGGAGGCTGATTTGGCTTGTGATGATGATGATGATGATG 2587
951 CCATACCTGATCTCTTCCACCTGCGCAATTAACCAAGTGAAGGAGGAAATTCACAGGGG 1020
2588 CCATACCTGATCTCTTCCACCTGCGCAATTAACCAAGTGAAGGAGGAAATTCACAGGGG 2647
1021 GCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
2648 GCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2707
1081 ACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
2708 ACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2767
1141 CTCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
2768 CTCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2827
1201 TGGTTTAAACGAGAAACATTTGAAAATGAGCAGATCTTTTCATGATGATGATGATGATG 1260
2828 TGGTTTAAACGAGAAACATTTGAAAATGAGCAGATCTTTTCATGATGATGATGATGATG 2887
1261 CTGTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
2888 CTGTTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2947
1321 TTCTCCACAGACTCCGCGCAGAGACACTAGTCTCATGAAAGAGTGTGCTGCTTGTCTTA 1380
2948 TTCTCCACAGACTCCGCGCAGAGACACTAGTCTCATGAAAGAGTGTGCTGCTTGTCTTA 3007
1381 ATATTCATATCAACACACACATTTCTGGAATTCATATTAATTAATGAGAGTGA 1440
3008 ATATTCATATCAACACACACATTTCTGGAATTCATATTAATTAATGAGAGTGA 3067
1441 TTAGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
3068 TTAGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3127

[illegible]

Query/Match	90.8%	Score 1528.4	DB 9	Length 2826	
Best Local Similarity	99.6%	Pred. No. 0			
Matches 1532	Conservative	0	Mismatches 6	Indels 0	Gaps 0
1 AACAAAGTGTGTCATCATCCACAGATGCTTTGGGGCCCTCTGCAGCTCAAGAACTAG 60					
D2 1281 AACAAAGTGTGTCATCATCCACAGATGCTTTGGGGCCCTCTGCAGCTCAAGAACTAG 1340					
QY 61 AGGAGCTGCCAAATGACAGAGGTTTACAGCATATCTCTCAAGTCAAGTTCAGAAACA 120					
D2 1341 AGGAGCTGCCAAATGACAGAGGTTTACAGCATATCTCTCAAGTCAAGTTCAGAAACA 1400					
QY 121 ATGGGCTCATATATGCTTTTGGGGCCCTTTCATCAGAAATGAGGTGTCTCCACGCT 180					
D2 1401 ATGGGCTCATATATGCTTTTGGGGCCCTTTCATCAGAAATGAGGTGTCTCTCACGCT 1460					
QY 181 CCATCAGCTTGAAGTAAGGAATTAACTCCAGACAGCCAGTGTGATGATGCCAG 240					
D2 1461 CCATCAGCTTGAAGTAAGGAATTAACTCCAGACAGCCAGTGTGATGATGATGCCAG 1520					
QY 241 TCATGTGACAGCACCCGTTGGGAAAGAGACACTTGTCTTATCAGCTGGACAAAGCAG 300					
D2 1521 TCATGTGACAGCACCCGTTGGGAAAGAGACACTTGTCTTATCAGCTGGACAAAGCAG 1580					
QY 301 CTCGCCAAATCTCTCTCTGGGATCCCACTGACAGACAGAAAGGTGGCTTTGATGAGACA 360					
D2 1581 CTCGCCAAATCTCTCTCTGGGATCCCACTGACAGACAGAAAGGTGGCTTTGATGAGACA 1640					
QY 361 AAAACACCAAAATGGCTCTCTCCCAATCCAGGCACTTCTAGTGTGGCACTTGGAAAT 420					
D2 1641 AAAACACCAAAATGGCTCTCTCCCAATCCAGGCACTTCTAGTGTGGCACTTGGAAAT 1700					
QY 421 ACAGTGTGACAGACGCTCAACAACTTGAACCTGTACGTCCCGTCCGTCCAAATG 480					
D2 1701 ACAGTGTGACAGACGCTCAACAACTTGAACCTGTACGTCCCGTCCGTCCAAATG 1760					
QY 481 CTACCTGCTCCAAATTACAGTACTTCCAAAACGAACAAAGACACACGCAATTTCCCA 540					
D2 1761 CTACCTGCTCCAAATTACAGTACTTCCAAAACGAACAAAGACACACGCAATTTCCCA 1820					
QY 541 GCGCTGTGAGTTTATGCAATATTTGCGCAAGGAGCTCCCAATTTCTAGGAGCCAGTG 600					
D2 1821 GCGCTGTGAGTTTATGCAATATTTGCGCAAGGAGCTCCCAATTTCTAGGAGCCAGTG 1880					
QY 601 TCACAGCCCTGATTGATCAGTAAATGAAAAACAGTACTTGGACATCTGGATATAG 660					
D2 1881 TCACAGCCCTGATTGATCAGTAAATGAAAAACAGTACTTGGACATCTGGATATAG 1940					
QY 661 GAGCAGTGTGATGCTTACTAAAGATGAGGTCTTAACTAAAGTATTTTACAACTATG 720					
D2 1941 GAGCAGTGTGATGCTTACTAAAGATGAGGTCTTAACTAAAGTATTTTACAACTATG 2000					
QY 721 ACACGAGTGTATATACAGTGTAAAGTGTGGGCTCTGGAGAGATTAAACGACCCAGAC 780					
D2 2001 ACACGAGTGTATATACAGTGTAAAGTGTGGGCTCTGGAGAGATTAAACGACCCAGAC 2060					
QY 781 GAGAGTGTATACCCAGACAGATGAGACACTGTACATCTGGCTGTGATTTGAGATATAG 840					
D2 2061 GAGAGTGTATACCCAGACAGATGAGACACTGTACATCTGGCTGTGATTTGAGATATAG 2120					
QY 841 AATATCAATGGAATCCACCAAGACCTGAATTTAATAGATGATGTTCAACACAGAGAG 900					
D2 2121 AATATCAATGGAATCCACCAAGACCTGAATTTAATAGATGATGTTCAACACAGAGAG 2180					
QY 901 TGTGTTTCACAGACATCTCTGGAGAGCTATTTGGCTTCGATGTCCCAATGCTC 960					
D2 2181 TGTGTTTCACAGACATCTCTGGAGAGCTATTTGGCTTCGATGTCCCAATGCTC 2240					

[illegible]

OY	1	AACAAAGTGGTGCATATCATCCACACAGTCCCTTTGGGGCCCTCTGCAGCTCAAGAACTAG	60
Db	1294	AGCAGAGGGGGGCCATCATCATCACTACAGTGGCCCTGGGACCGGCTGAAAGACTTG	135
OY	61	AGGAGCTGTCCAAAATGACAGAGGCTTTACAGACATATGCTTCAGATCAAGTTTCAGAAC	120
Db	1354	AGCAGCTGTCCAAAATGACAGAGGCGCTGCAGACATATCTCTCGATCAGGTTTCAGAA	141
OY	121	ATGGCTCATATGATGCTTTTGGGGCCCTTTTCATCAGAAATGGAGCTGTCTCTCAGCGT	180
Db	1414	ATGGCTCTTGTTGATGCTTTTCCAGCACTCTCTCTCAGAAATGGCGGATCGCTCAGCACT	1473
OY	181	CCATCCAGCTTTGAGAGTAAAGGATTTAACCTCCAGAAACAGCCAGTGTAGTAATGGCAG	240
Db	1474	CCATCCAGCTTGAAGAGCAGGAGGAGTATATCTCCAGATTAACCAATGATGTAATGGCTAG	1533
OY	241	TGATGTGGAGACAGCAGCGTGGGAAAGACCTTTGTTCTATCACTCTGGACACAGCAG	300
Db	1534	TGATGTGGAGAGCTGTGGTGGGACAGGACACCTTTGTTCTATCACTCTGGACACAGCATC	1593
OY	301	CTCCCAAAATCTCTCTGGGATCCCAAGTGGACAGAAAGAGGAGGCTTTGTATGGACA	360
Db	1594	CTCCCAAAATTTATCTGGGATCCCAAGGAGGATGGAAACAAAATGGTTTATATCAAGACA	1655
OY	361	AAAACACCAAAATGGCTTACCCTCCAAATCCAGGCAATGCTAAGTTGGCACTTGGAAAT	420
Db	1654	CAACCACTAAGGTGGCTTACTCTCCAAAGTCCCAAGCAGCGCTAAGGTTGGCTTTGGAAAT	1713
OY	421	ACAGTCTCAAGCAAGCTCACAAAACCTCTGACCTGATGCACAGTCCCGTGGGTCCATG	480
Db	1714	ACAGATTCAGGCAAGCTCACAGACTCTCACTTGTACTGTCACTCTCCGTGCACAGAACTG	1773
OY	481	CTACCTGTCTCAATTAACAGTACTTCCAAAAGAACAGACACACCAAAATTTCCCA	540
Db	1774	CTACACTGTCTCTATTAACAGTGAACCCCGGTAGTGAATTAAGAACACAGGAAATTTCCCA	1833
OY	541	GCCCTGTGGTATTTATGCAAAATATTCGGCAAGAGCCTCCCAATTCACAGGGCAGTG	600
Db	1834	GCCCTGTACAGTGTATGCAAGCAATTCGCCAAGAGAGCTTGCCCTATTTCTCAAGGGCAGG	1893
OY	601	TCACAGCCCTGATTAATCAGTGAATGGAAGAAAACATTACTTGAATCACTAGTGAATATG	660
Db	1894	TCACAGCCTTATTAATGATGTGTAATGGAAGAAAACAGTAACCTCGAATTAAGTAATAG	1953
OY	661	GAGCAGGTGTGATGCTACTAAGAGTAGGGGTCTCTACTCMAAGTATTTCACAACTATG	720
Db	1954	GAGCAGGTGCGGATGCGCACCAAGAAATGATGTCTACTCAAGGTTTTCACAGGTTTTG	2013
OY	721	ACACGAATGTGATGATACAGTGTAAAGTGGGGCTCTGGGAGAGATTAACGACGCACAC	780
Db	2014	ATGCAAAATGGTATGATACAGGTTTAAATATATGGGCTCTGGGAGAGTCACTTACAGCAAC	2073
OY	781	GGAGAGTGAATACCCAGCAGAGTGGAGCACTGTACTATACCTGGCTGGATTAAGAAATG	840
Db	2074	AGAGAGCAGCACTCTCCAGAACAGAGAGCCATGTACTATAGATGGCTGATTTGAGGATGGT	2133
OY	841	AAATATCATGGAATCCACCAAGAGCTGAATTTAATTAAGATGATGTGTAACCAAGCAGAG	900
Db	2134	AAGTATGAATACCCACAGCTCTGAATCTA-----GTTATGTCTCAAGCAAGCAGAGC	2187
OY	901	TGTGTTTCAGCAACAATCTCTCGGAGAGCTCATTTGTGGCTTCTGATGTC---CAAAATG	957
Db	2188	TGTGTTTCAGCAAGCAATCTTCACAGGGGAGTCTTTGTGGCCACAAATGCCCCGACAG	2247
OY	958	CTCCCAATACCTGATCTTTCCCACTGTGGCAATTCACGCACTGAAGGGGGAATTCAG	1017
Db	2248	CTCCCAATTCCTGACTCTCTTTCACACCTGTCAAAATCACTACCTGAAGGGCAGCACTCAG	2307
OY	1018	GGGCGAGCTCATATTAATCTACTATTTGGACAGCTCTCTGGGATGTTATGACATGGAGAC	1077
Db	2308	GGCAGAACTGTGTAATCTCTACGTGGAGGGCTCTCTGGGATGATCTACGACCAAGGAGAG	2367
OY	1078	CTCAAGATATATCATTCGAATGAATACAGTATTTCTGATCTCAGAGCAAGATTTCAATG	1137

DB	1138	AAATCTTTAAAGTGAATAGTACTGCTCTCATCCCAAAGAGCAACCTTGAGGAAGTCT	1197
DB	2428	CCTCACTCCCAAGTGAACACCTACCGGCTTATCCCAAGAGCCAGCTGTGAGGAATCT	2487
QY	1198	TTTTTTTAAACAGAAAACATTACTTTTGAAGAAAGCCAGATCTTTTCATTGCTATTC	1257
DB	2486	TTGAGTTTAACTGGGAGGCAACCTTTTGGAAATGGCAGATATCTTCAATGGTATCC	2547
QY	1258	AGCGTGTGATTAAGGTGCATCTGAATATCAAGAAATATCCAACTTGCAGAGATCTTTTGT	1317
DB	2548	AGCGTGTGATTAAGTGCATCTGAATATCAAGAAATATCCAACTTGCAGAGAGTCTGTCT	2607
QY	1318	TTATTTCTCCACAGACTCCGCCAGAGACACCTAGTCTGATGAATAAAGTCTGCTCTGTC	1377
DB	2638	TCATCCCGCCCTCAG-----GAGCCGCCCATTCCTCCCAAGACATCACTCCCTTGTGC	2658
QY	1378	CTAATATTCATTTCAACAGACACCATTCCTCGGCAATTCATCTTTTAAATATATGGAAGT	1437
DB	2658	CTGACATTCAGCATTCACACACCATTCCTCGGCAATTCACAGTGTGAAGATTAATGGAAGT	2718
QY	1438	GGATAGAGAACTGCAGCTGTCAATAGCCCTAGCGCTGAATTTTGTGCAGATTAATTAAT	1497
DB	2718	GGCTAGGGAATATGACAGTGACATGATGTTTGCATGAATTTTTCAGGCAAGAAATCAACC	2778
QY	1498	AAATCATTCCTCTTTTGTGATTAATAATTTCTAAATATGATTTTATGACTCTCTCT	1557
DB	2779	A-----GTCAATTCCTTTCACCTGGAGATTTTCTAAATAATGTAATTAAGCTTCTCTCT	2830
QY	1558	AGGGGAGCATATATCTAA 1574	
DB	2831	AGGGGAGCATATATCTAA 2847	
RESULT	7		
LOCUS	195746	878 bp	DNA
DEFINITION	Sequence 8 from patent US 5733748.		PAT
ACCESSION	195746		01-DEC-1998
VERSION	195746.1	GI:3940216	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 878)		
AUTHORS	Ku, G. and Rosen, C.		
TITLE	Colon specific genes and proteins		
JOURNAL	Patent: US 5733748-A 8 31-MAR-1998;		
FEATURES	Location/Qualifiers		
Source	1..878		
BASE COUNT	257 a 179 c 188 g 241 t	13 others	
CRIGIN			
Query Match	47.0%: Score 790.8: DB 6: Length 878:		
Best Local Similarity	97.9%: Pred. No. 2.5e-190:		
Matches	820: Conservative 1: Mismatches 13: Indels 2: Gaps 2:		
QY	632	TGCTACTCAAGAGTATTTACAACTTATTCACAGCAATGATAGATACAGTGTAAAGTGG	751
DB	1	TTCTACTCAAGAGTATTTACAACTTATTCACAGCAATGATAGATACAGTGTAAAGTGG	60
QY	752	GGCTTGGGAGAGTTAAACGACGACGAGAGAGTGAATCCCGACAGAGTGGAGCACT	811
DB	51	GGCTTGGGAGAGTTAAACGACGACGAGAGAGTGAATCCCGACAGAGTGGAGCACT	120
QY	832	GTACATTAACCTGGCTGGATTTGAGATGATGAATTAACATGGAATTCACCAAGACCTGAAT	871
DB	121	GTACATTAACCTGGCTGGATTTGAGATGATGAATTAACATGGAATTCACCAAGACCTGAAT	180

QY 872 TAATAAGATGATGTCACACAGCAAGAGTGTCTTCAGCAGAACTCTCGGGAGGCTC 931
DB 181 TAATAAGATGATGTCACACAGCAAGAGTGTCTTCAGCAGAACTCTCGGGAGGCTC 240
QY 932 ATTGGGCTCTGATGTCACCAATGCTCCATACCTGATCTCTCCAGCTGCGCAAT 991
DB 241 ATTGGGCTCTGATGTCACCAATGCTCCATACCTGATCTCTCCAGCTGCGCAAT 300
QY 992 CACCGACCTGAAAGCGGAAATTCACGGGCGAGCTCTCATTAATCTGATGAGACGCTCC 1051
DB 301 CACCGACCTGAAAGCGGAAATTCACGGGCGAGCTCTCATTAATCTGATGAGACGCTCC 360
QY 1052 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1111
DB 361 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 1112 TCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1171
DB 421 TCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1172 AAGGAGGCACTCTGAGGAGCTCTGATGATGATGATGATGATGATGATGATGAT 1231
DB 481 AAGGAGGCACTCTGAGGAGCTCTGATGATGATGATGATGATGATGATGATGAT 540
QY 1232 TGGGACAGATCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
DB 541 TGGGACAGATCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 1292 ATCCAACTGTCAGAGATCTTTTATGATGATGATGATGATGATGATGATGATGAT 1351
DB 601 ATCCAACTGTCAGAGATCTTTTATGATGATGATGATGATGATGATGATGATGAT 660
QY 1352 TCCGTATGAAAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1410
DB 661 TCCGTATGAAAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
QY 1411 TTCACATTTTAAAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1470
DB 721 TTCACATTTTAAAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1471 GGTGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1528
DB 781 GGTGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837

RESULT 8
AF127035 3204 bp mRNA PRI 11-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 2 (CAC2)
DEFINITION
ACCESSION AF127035
VERSION AF127035.1 GI:5726288
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3204)
AUTHORS Agnel, M., Vermaut, J., and Culoussou, J. M.
TITLE Identification of three novel members of the calcium-dependent
chloride channel (CaCC) family predominantly expressed in the
digestive tract and trachea
JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792
REFERENCE 2 (bases 1 to 3204)
AUTHORS Agnel, M., and Culoussou, J. M.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des
Carrières, Rueil-Malmaison 92500, France
FEATURES
Source 1.3204
/organism="Homo sapiens"

/db_xref="taxon:9606"
/issue_type="colon"
1..3204
/gene="CAC2"
29..2782
/gene="CAC2"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 2"
/protein_id="AA048398.1"
/db_xref="GI:5726289"
/translation="MGLEFGEFLVLLHOSNTSFIKNNGEFIVIDPSVE
DEKIDIDDMWTASTYLEFEATKEFFKNYSILPEMKENPQYRKHNHKNAD
VIAPEPTLGRDEPYKOTFECEKEEYHIFEDLLDKKONGYGPGLFVEMHL
RMVPEYEDNDDEPYKSKLEATCSGSIERNVYCGKSSRCRDIESTTKL
YKDDQFPDPKOTKESIMFMSIDSVFECNEKHNDPESLQIKNFRSTWTVI
SNSEDKNTIPWTPPPPPVPSLIRQIVLCVLYDKSSMGCKPLNRMNKAHFL
LQIVENGSVNWHEDSTATVYKLIQISDSERNLMLGLPPYPLGYSISGIFYA
FOYIGSHFVSDEAONNGLIDAFGALTSNLTDSKLOLESGLNLSNANDVI
KITGSHFVSDEAONNGLIDAFGALTSNLTDSKLOLESGLNLSNANDVI
IDSVKQKDFEFLITNSLPSISIMPSGIMENFVDAKSKMAYLISGTAKVGTVA
VNOAKANPEFLITVTSRAANSVPPITVYNNKMNVDSPSPHIVAYEIIQGYPV
LCANVAFIESONGHTFVLELDNGAGSFKNDGVSRFTAYTENGRYSLKVRAG
GANTARLKLRLPPLNRAAYIPGVNNGEILANPPEPIDEDYTTLEDPSRISGAFV
VSQVPSLPDPPSOITDLDATVHEDIIITLWAPGNDFGVQYRIYIRISAIL
DLRSEFDALQVYTTDLSPKANSKESFKEENISEENATHFIFAIKIDRSNLTSK
VSLIAQVTLFIPQANDPIDPPPTPPPPDKSHNSGVNISTLVISGVIVYNFLIS
TII"
BASE COUNT 1098 a 594 c 633 g 879 t
ORIGIN
Query Match 40.0%; Score 673.8; DB 9; Length 3204;
Best local similarity 69.2%; Pred. No. 1.6e-160;
Matches 966; Conservative 0; Mismatches 417; Indels 12; Gaps 3;
QY 1 AACAAAGTGTGCCATCATCATCAGACAGTCTGTTGGGCGCTCTGACGTCAAGAACTAG 60
DB 1308 AACAAAGTGTGCCATCATCATCAGACAGTCTGTTGGGCGCTCTGACGTCAAGAACTAG 1367
QY 61 AGNAGCTGTCCAAATGACAGAGAGCTTTACAGACATATGCTTACATGATCAAGTTCAGACA 120
DB 1358 TAGAGATGACCAAGATATACAGAGAGAGAGTATTTTATGTTTACAGAGAGTTCAGANCA 1427
QY 121 ATGGCCTATGATGCTTTTGGGGCCCTTTCATCAGAAATGAGAGCTGTCTCAGGCT 180
DB 1428 ATGGCCTATGATGCTTTTGGGGCCCTTTCATCAGAAATGAGAGCTGTCTCAGGAGT 1487
QY 131 CCATTCAGCTTGAAGATGAAGGATTAACCTCCAGAACAGCCAGTGGATGATGACAG 240
DB 1488 CCCTTCAGCTGGAAGATGAAGGATTAACCTGAATAGTATGCTGTGATACAGACAGCTG 1547
QY 241 TGAATGTCGACAGCACCCTGGGAAAGACACTTTGTTTCTTATACCTGAGCAACAGCAGC 300
DB 1548 TCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1607
QY 301 CTGCCCAAAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 1608 CTGCCCAAAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 367
QY 361 AAAACACCAAAATGAGCTTACCTCCAAATCCAGAGCATTTGATGATGATGATGATGAT 420
DB 1668 CAATCTTCAAAATGAGCTTACCTCCAAATCCAGAGCATTTGATGATGATGATGATGAT 427
QY 421 ACAAGTCTG-----AAGCAGCTCACAAACCTTGACCTGACGTGACCTCCGTCGCT 474
DB 1728 ACAATCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1787
QY 485 CCAATGCTACCTGCTCCCAATTAAGTGCATTCACAAATGAGCAAGCAAGCAAGCAAGCA 534
DB 1788 CCAATGCTGCTGCTCCCAATTAAGTGCATTCACAAATGAGCAAGCAAGCAAGCAAG 1847
QY 535 TCCCAAGCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 594

Db 1848 TCCACGAGCCATGATGTTTACGAGAAATCTACAGAGATGATGATCTGTTCTTGAG 1907
 Oy 595 CCAAGTGTACAGCCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 654
 Db 1908 CCAATGTGACGCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATG 1967
 Oy 655 ATATGAGAGAGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 714
 Db 1968 ATATGAGAGAGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2027
 Oy 715 CTTATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 774
 Db 2028 CATATACAGAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2087
 Oy 775 CCAAGAGAGAGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 834
 Db 2088 CCAAGAGAGAGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2147
 Oy 835 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 894
 Db 2148 ACGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2204
 Oy 895 AGCAGAGAGAGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 954
 Db 2205 CTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2264
 Oy 955 ATGCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
 Db 2265 GCTTGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2324
 Oy 1015 ACGGAGAGAGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1074
 Db 2325 ATGAGAGAGAGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2381
 Oy 1075 CAGCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1134
 Db 2382 AAGTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2441
 Oy 1135 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1194
 Db 2442 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2501
 Oy 1195 TCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1254
 Db 2502 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2561
 Oy 1255 TTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1314
 Db 2562 TTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2621
 Oy 1315 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1374
 Db 2622 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2681
 Oy 1375 GTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1389
 Db 2682 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2696

RESULT 9
 AK000138 1895 bp mRNA PRI 22-FEB-2000
 LOCUS Homo sapiens cDNA FLJ20131 fls, clone COL06357.
 DEFINITION AK000138
 ACCESSION AK000138
 VERSION AK000138.1 GI:7020030
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens colon cDNA to mRNA, clone_11b.COL clone:COL06357.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (sites)
 AUTHORS Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y.,

Okamoto,S., Ohtani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
 Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 Unpublished (2000)
 2 (bases 1 to 1895)
 REFERENCES
 Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
 Shibahara,T., Tanaka,T. and Nakamura,Y.
 Direct Submission
 Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases, Sumio
 Sugano, Institute of Medical Science, University of Tokyo, Department
 of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan: cDNA full insert
 sequencing: Research Association for Biotechnology: cDNA library
 construction, 5'- & 3'-end one pass sequencing: Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).
 FEATURES
 source
 location/Qualifiers
 1..1895
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COL06357"
 /clone_11b="COL"
 /issue_type="colon"
 /note="cloning vector pME18SFL3"
 66..1475
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="PAA90969.1"
 /db_xref="GI:7020031"
 /translation="MSKTTGSHHYVDEQNGLIDAFGLTSGNTDISKLSQLES
 KGLTNSNAMNDVILDSYGVKDFEFLITWNSLPSTISLDSGTLMENTVDATSK
 MAYLISPGTAKVGMWNLQAKAPETLITTVTSRAANSVPLITVAKNKNVDNSFP
 SPMLVAVELILOGVPLVAVANTATIEBSONGTEVLELDNAGADSKRNQVSRFT
 AYTFNGHRSKLKVRHAGANFARLRLPLNSAAIYIPMWVGEIEANPPEDEIDED
 TTLEDFRSRTSGSGFVYVSOVSESLPLPOYRPLSDTIDAVYHEKTLITWPAEDND
 VGRVRYRTIRSAIIDLRSDFDALVNTDLSPKRNSSESAFPEPENSFNATH
 IFVAKSIDKSNLSKYSNINQVILFIPQANPDIDPPTPTPPDPSHNSGVNISTL
 VLSVIGSVVIVNLFSTTI"
 BASE COUNT 658 a 352 c 357 g 528 t
 ORIGIN
 Query Match 39.9%; Score 672.2; DB 9; Length 1895;
 Best Local Similarity 69.2%; Pred. No. 3.8e-160; Mismatches 418; Indels 12; Gaps 3;
 Matches 965; Conservative 0;
 Db 1 AACAAAGTGGGGCCATTGTTCAATTTATGCTTGGGAAGAGCTGCTGATGAGAGTGA 60
 Oy 1 AACAAAGTGGGGCCATTGTTCAATTTATGCTTGGGAAGAGCTGCTGATGAGAGTGA 60
 Db 51 AGGAGCTGTCCAAATGACAGAGAGTTTACAGACATATGCTTACAGATCAAGTTCAAGA 120
 Oy 51 AGGAGCTGTCCAAATGACAGAGAGTTTACAGACATATGCTTACAGATCAAGTTCAAGA 120
 Db 61 TAGAGATGAGCAAGATTAACAGAGAGATTTATGTTTACAGATGAGAGTCAAGACA 120
 Oy 61 TAGAGATGAGCAAGATTAACAGAGAGATTTATGTTTACAGATGAGAGTCAAGACA 120
 Db 121 ATGGCTCATGATGCTTTTGGGCGCTTTCATGAGGAATGAGAGCTGCTCTCAGCGCT 180
 Oy 121 ATGGCTCATGATGCTTTTGGGCGCTTTCATGAGGAATGAGAGCTGCTCTCAGCGCT 180
 Db 121 ATGGCTCATGATGCTTTTGGGCGCTTTCATGAGGAATGAGAGCTGCTCTCAGCGCT 180
 Oy 121 ATGGCTCATGATGCTTTTGGGCGCTTTCATGAGGAATGAGAGCTGCTCTCAGCGCT 180
 Db 141 CCATCAGCTGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 Oy 141 CCATCAGCTGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 Db 181 CCGTTCAGCTGCAAGATTAAGGATTAACACATGATGATGATGATGATGATGATGATGATG 240
 Oy 181 CCGTTCAGCTGCAAGATTAAGGATTAACACATGATGATGATGATGATGATGATGATGATG 240
 Db 241 TCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 Oy 241 TCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 Db 301 CTCGCCAATATCTTCTCTGGGATCCAGTGGACAGACAGAGAGTGGCTTTGATGAGACA 360
 Oy 301 CTCGCCAATATCTTCTCTGGGATCCAGTGGACAGACAGAGAGTGGCTTTGATGAGACA 360
 Db 301 CTCGCCAATATCTTCTCTGGGATCCAGTGGACAGACAGAGATGAGAAATGAGAAATTTCA 360

QY	361	AAAACACAAAATGGCCTACCTACCTCAATCCCAAGCAATTCCTAAGGTGGTGGACCTTGGAAAT	420
Db	361	CAACTTCACAAAATGGCCTATCTCAGTATTCACAGAACTCCAAAGGTGGGCACTTGGGCAT	420
QY	421	ACAGCTGC-----AAGCAAGCTCACAAACCTTGAACCTGACTGTACGTCCTCGTGCST	474
Db	421	ACAAATCTTCAAGCCAAAGCCAAAGCCAGAAAACATTAATCTTAACAGTAATCTTCGAGCAG	480
QY	475	CCAAATGCTACCCCTGCCCAATTAACAGTGCCTTCCAAAAGCAAGGACACACAGAAAT	534
Db	481	CAAAATCTTCTGCTCCCTCCCAATACAGTGAATGCTAAATGAATTAAGACGTAAACAGTT	540
QY	535	TCCCCAGCCCTCTGTGATTAATGCAAAATATTCGCCAAGAGGAGCCCTCCCAATTCACAGG	594
Db	541	TCGCCAGCCCAATGATTTGTTAAGCAGAAATTCACAGAAATATCTACAGATATCTACCTGTTCTGCAG	600
QY	595	CCAGTGTACAGCCCTTAATGTAATCACTGAATGGAAGAAAACAGTTACCTTGAACACTACG	654
Db	601	CCAAATGACAGCTTTTCAATTAATACAGAAATGAGCAATACAGAAAGTTTGTGAACCTTTGG	660
QY	655	ATAATGGAGCGGTGATGATGCTACATPAGAGATACGGGTGTCTACCAAGTATTCACAA	714
Db	661	ATAATGGAGCGGCGTGTATCTTCAAGATATGGAAGTACTACAGATATTTACAG	720
QY	715	CTTATGACACGAATGATGATACAGTGTAAAGATGGCGGCTCTGGAGAGAGTTAAACGAG	774
Db	721	CATATACAGAAATATGGCAGATATAGCTTAAAGTTCGGGCTCATGAGAGGCAAAACATG	780
QY	775	CCAAACGAGAGATATACCCACAGAGTGGAGCACTGTACATACCTGCTGGATTTGAGA	834
Db	781	CCAGGCTPAAATTAAGGCGCTCCACTGTAATGAGACCGGCTGATACATACCAAGCTGGATGTA	840
QY	835	ATGATGAATATCAATGGAATCCACCAAGACCTGAATTAATAGATGATGTTTCAACACA	894
Db	841	ACGGGGAATATGAAGCAACCCGCGCAAGACTTAATTAATGAT---GAGGATACACAGACA	897
QY	895	AGCAATGTGTTTCAGCAGAAACATCCTCGGAGGCTCATATTGTGCTCTGATGTCCAA	954
Db	898	CCTTGGAGGATTTTCAGCCGAACAGCAATCCGAGAGTGCATTTGGTATTCACAAAGTCCAA	957
QY	955	ATGTCGCCATACCGATCTCTCCACACTGGCCAAATACCGCACTGAAGGGGAAATTC	1011
Db	958	GCCCTCCCTGCTGACCAATACCCACCAAGTCAAAATCACAGACTTGAATGCCACAGTTC	1011
QY	1015	ACGGGGCGAGTCTATTAAATCTGACTTGGACAGCTCCTCGGAGATGATATGACCAATGAA	1071
Db	1018	ATGAGG---ATPAGATATATCTTACATGAGACACACAGAGATAAATTTTGATGTTGAA	1071
QY	1075	CAGTCCACAGTATATCTATTCGATTAAGTACAAAGTATTTCTGATCTCAAGACAGTCTA	1133
Db	1075	AAGTTCAACGTTATATATCAAGAAATTAAGTCAAGTATCTTGGATCTPAAAGACAGTTTGG	1133
QY	1135	ATGATATCTCTCAAGTCAAAATCTACTGCTCTCAATCCCAAGAAAGCAACTGAGAGAG	1199
Db	1135	ACGATGCTCTCAAGTCAAAATCTACTGATCTGTCTACCAAGAGAGGCACTCCAAAGGAA	1199
QY	1195	TCCTTTTGTGTTAAACCGAAAACATTAATCTTTGAAATGGCAGACAGATCTTTTCATTTGCTA	1251
Db	1195	GCTTTGCAATTTAAACCGAAAATATCTCAAGAAAATTAAGCAACCCCAATATTTATGCGCA	1251
QY	1255	TTCAAGGCGGTTGATPAGGTGATGTGAATACAGAAATATCAACATTTGACGAGTATCTT	1311
Db	1255	TTTAAAGATATGATPAAAGCAATTTGACATCAAAAGTATCAACATTTGCCAAGATACCTT	1311
QY	1315	TGTTTATTTCTCCACAGACTCTCCGACAGACACCTAGTCTGATGAAGAGTCTGCTCTT	1371
Db	1315	TGTTTATCTCCCAAGCAAAATCTGATGACATGATCTCTACTCTACTCTACTCTACTCTC	1371
QY	1375	GTCCCAATATTCATA	1389
Db	1375	CTGATTAAGTCTATA	1389

LOCUS	AK000072	3221 bp	mRNA	PR1	22-FEB-2000
DEFINITION	Homo sapiens CDNA FLJ20065 fis, clone COL01613, highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.				
ACCESSION	AK000072				
VERSION	AK000072.1 GI:7019922				
KEYWORDS	oligo:capping; fis (full insert sequence).				
SOURCE	Homo sapiens colon CDNA to mRNA, clone_lib:COL clone:COL01613.				
ORGANISM	Homo sapiens				
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Kawabata,A., Hikiji,T., Kobayake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.				
TITLE	NEBO human cDNA sequencing project				
JOURNAL	Unpublished (2000)				
REFERENCE	2 (bases 1 to 3221)				
AUTHORS	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ems.u-tokyo.ac.jp, Tel:81-3-5449-5266, Fax:81-3-5449-5416)				
COMMENT	NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; CDNA full insert sequencing: Research Association for Biotechnology; CDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).				
FEATURES	<p>Location/Qualifiers</p> <p>1..3221</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="COL01613"</p> <p>/clone_lib="COL"</p> <p>/tissue_type="colon"</p> <p>/note="cloning vector pME18SFL3"</p> <p>1..3221</p> <p>/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"</p>				
BASE COUNT	1105 a 600 c 634 g 882 t				
ORIGIN					
Query/Match	39.7%; Score 667.6; DB 9; Length 3221;				
Best Local Similarity	69.5%; Pred. No. 5.9e-159;				
Matches	968; Conservative 0; Mismatches 409; Indels 15; Gaps 4.				
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	6	6	6	6
7	7	7	7	7	7
8	8	8	8	8	8
9	9	9	9	9	9
10	10	10	10	10	10
11	11	11	11	11	11
12	12	12	12	12	12
13	13	13	13	13	13
14	14	14	14	14	14
15	15	15	15	15	15
16	16	16	16	16	16
17	17	17	17	17	17
18	18	18	18	18	18
19	19	19	19	19	19
20	20	20	20	20	20
21	21	21	21	21	21
22	22	22	22	22	22
23	23	23	23	23	23
24	24	24	24	24	24
25	25	25	25	25	25
26	26	26	26	26	26
27	27	27	27	27	27
28	28	28	28	28	28
29	29	29	29	29	29
30	30	30	30	30	30
31	31	31	31	31	31
32	32	32	32	32	32
33	33	33	33	33	33
34	34	34	34	34	34
35	35	35	35	35	35
36	36	36	36	36	36
37	37	37	37	37	37
38	38	38	38	38	38
39	39	39	39	39	39
40	40	40	40	40	40
41	41	41	41	41	41
42	42	42	42	42	42
43	43	43	43	43	43
44	44	44	44	44	44
45	45	45	45	45	45
46	46	46	46	46	46
47	47	47	47	47	47
48	48	48	48	48	48
49	49	49	49	49	49
50	50	50	50	50	50
51	51	51	51	51	51
52	52	52	52	52	52
53					

QY 301 CTCGCCAAATCTCTCTGGATCCAGTGGACAGAACAGTGGCTTGTAGTGACA 360
 Db 1622 CTCGCCAGATATTCCTGTGGATCCAGTGGACAGAACAGTGGCTTGTAGTGACA 1681
 QY 361 AAAACACCAAAATGGCTCTACCTCCAAATCCAGCATTTGCTAGGTGGACCTTGAAT 420
 Db 1682 CAATCTTCAAAATGGCTCTACCTCCAGCATTTGCTAGGTGGACCTTGAAT 1741
 QY 421 ACAGCTGTC-----AAGCAAGCTCACAAACCTTGACCTGCTGACGTCCTGCGT 474
 Db 1742 ACAATCTTCAAGCCAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1801
 QY 475 CCAATGCTACCTGCTCCAAATACAGTCTCCAAACAGAACAGCAACCAAGCAAT 534
 Db 1802 CAAATCTTCTGTGCTCCAAATACAGTCTCCAAACAGAACAGCAACCAAGCAAT 1861
 QY 535 TCCCAAGCCCTCTGCTATTTATGCAATATTTGCCAAGGACCTCCCAATTTCTCAGG 594
 Db 1862 TCCCAAGCCCAATGCTATTTATGCAAGTCTTACCAAGGATGCTGCTGCTTGGAG 1921
 QY 595 CCAGTGTACAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 654
 Db 1922 CCAATGTACCTGCTTCTATTTATGCAAGTGTGATGATGATGATGATGATGATGAT 1981
 QY 655 ATAAATGAGCAGGTGCTGATGCTACTAAGATGAGGATGAGGATGAGGATGAGGAT 714
 Db 1982 ATAAATGAGCAGGTGCTGATGCTACTAAGATGAGGATGAGGATGAGGATGAGGAT 2041
 QY 715 CTATGACACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 774
 Db 2042 CATATACGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2101
 QY 775 CCAGACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
 Db 2102 CCAGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2161
 QY 835 ATGATGAATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894
 Db 2162 ACGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2218
 QY 895 AGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 954
 Db 2219 CTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2278
 QY 955 ATGCTCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014
 Db 2279 GCTTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2338
 QY 1015 ACGGAGGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1074
 Db 2339 ATGAGG---ATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2395
 QY 1075 CAGTCAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134
 Db 2396 AAGTCAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2455
 QY 1135 ATGATCTCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
 Db 2456 ATGATCTCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2515
 QY 1195 TCTTTTGTGTTAAACAGAAATACATTTGTAATGSCACAGATCTTTTCAATGCTA 1254
 Db 2516 GCTTTGATTTAAACAGAAATATCTCAGAAATGCAACCACTATTTATTTGCA 2575
 QY 1255 TTGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
 Db 2576 TTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2635
 QY 1315 TGTATTATCTCCAGACACTCCGACAGAG---CACCTAGTCTGATGATGATGATGATG 1371
 Db 2636 TGTATTATCTCCAGACAAATCTGATGATGATGATGATGATGATGATGATGATGATG 2695

QY 1372 CTGTCTCTAATA 1383
 Db 2696 CTACTCTTGATA 2707
 RESULT 1311
 AX092338 3265 bp DNA PAT 21-MAR-2001
 LOCUS AX092338
 DEFINITION Sequence 69 from Patent WO0116318.
 ACCESSION AX092338
 VERSION AX092338.1 GI:13444483
 KEYWORDS
 SOURCE human;
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3265)
 Godowski, P.J., Grimaldi, C.J., Gurney, A.L., Watanabe, C.K. and
 Wood, W.I.
 TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
 the same
 JOURNAL Patent: WO 0116318-A 69 08-MAR-2001;
 Genentech, Inc. (US)
 FEATURES
 source location/Qualifiers
 1..3265
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 1159 a 596 c 632 g 878 t
 ORIGIN
 Query Match 39.7%; Score 667.6; DB 6; Length 3265;
 Best Local Similarity 69.5%; Pred. No. 5; 9e-159;
 Matches: 968; Conservative 0; Mismatches 409; Indels 15; Gaps 4;
 QY 1 AACAAAGTGTGCTCATCTCCACACAGTGGCTTGGGGCCCTGTCAGCTCAGACTAG 60
 Db 1364 AACAAAGTGTGCTCATCTCCACACAGTGGCTTGGGGCCCTGTCAGACTAG 1363
 QY 61 AGGAGCTGTCCAAATATGACAGAGGATTTACAGACATATGCTTACATGATGATGATG 120
 Db 1364 TAGAGATGAGCAAGATTAACAGAGAGAGTCAATTTTATGTTTCAAGATGATGATG 1423
 QY 121 ATGGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 Db 1424 ATGGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1483
 QY 191 CCATCCAGCTTGAGAGTGAAGGATTAACCTTCAGAACAGCAGTGAATGATGATGATG 240
 Db 1484 CCCTTCAGCTGGAAGTGAAGGATTAACCTTCAGAACAGCAGTGAATGATGATGATG 1543
 QY 241 TGATGTGAGACAGCCTGGAAGGATGATGATGATGATGATGATGATGATGATGATG 300
 Db 1544 TCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1603
 QY 301 CTCGCCAAATCTCTCTGGAGTCCAGTGGACAGAACAGTGGCTTGTAGTGACA 360
 Db 1604 CTCGCCAGATTTCTCTGGAGTCCAGTGGACAGAACAGTGGCTTGTAGTGACA 1663
 QY 361 AAAACACCAAAATGGCTCTACCTCCAAATCCAGCATTTGCTAGGTGGACCTTGAAT 420
 Db 1664 CAATCTTCAAAATGGCTCTACCTCCAGCATTTGCTAGGTGGACCTTGAAT 1723
 QY 421 ACAGCTGTC-----AAGCAAGCTCACAAACCTTGACCTGCTGACGTCCTGCGT 474
 Db 1724 ACAATCTTCAAGCCAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1783
 QY 475 CCAATGCTACCTGCTCCAAATACAGTCTCCAAACAGAACAGCAACCAAGCAAT 534
 Db 1784 CAAATCTTCTGTGCTCCAAATACAGTCTCCAAACAGAACAGCAACCAAGCAAT 1843
 QY 535 TCCCAAGCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 594


```
|||||
Db 32982 GTGATAGGACAACTGCGCTCAATGAGCTAGGCGGATTTTGTCAATATAA 33041
OY 1496 ATAAATATTCATCTTTTGTGATTAATAATTTCTTAATATGATTTAGACTTCT 1555
Db 33042 ATAAATATTCATCTTTTGTGATTAATAATTTCTTAATATGATTTAGACTTCT 33101
OY 1556 GTAGGGGGCGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1615
Db 33102 GTAGGGGGCGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 33161
OY 1616 GCGATATCTAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1675
Db 33162 GCGATATCTAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 33221
OY 1676 ACTGGGTA 1683
|||||
Db 33222 ACTGGGTA 33229

RESULT 13
AL358950 164891 bp DNA HTG 23-JAN-2001
LOCUS Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN
DEFINITION PROGRAMS *** 10 unordered pieces.
ACCESSION AL358950
VERSION AL358950.4 GI:12539689
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Plumb.B.
1 (bases 1 to 164891)
Direct Submission
Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Jan 26, 2001 this sequence version replaced gi:9988471.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba444C12
----- Summary Statistics
Sequencing program: XCAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least Q40
Consensus quality: 162610 bases at least Q30
Consensus quality: 163328 bases at least Q20
Insert size: 163891; sum-of-contigs
Insert size: 163373; 6.6% error; agarose-fp
Quality coverage: 5.19x in Q20 bases; sum-of-contigs quality
coverage: 5.21x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 34236: contig of 34236 bp in length
* 34237 34336: gap of 100 bp
* 34337 47372: contig of 13036 bp in length
* 47373 47472: gap of 100 bp
* 47473 67025: contig of 19553 bp in length
* 67026 67125: gap of 100 bp
* 67126 71239: contig of 4114 bp in length
```

```
71240 71339: gap of 100 bp
* 71340 83903: contig of 12564 bp in length
* 83904 84003: gap of 100 bp
* 84004 91357: contig of 7354 bp in length
* 91358 91457: gap of 100 bp
* 91458 96452: contig of 4995 bp in length
* 96453 96552: gap of 100 bp
* 96553 148189: contig of 51637 bp in length
* 148190 148289: gap of 100 bp
* 148290 159778: contig of 11489 bp in length
* 159779 159878: gap of 100 bp
* 159879 164891: contig of 5013 bp in length.
FEATURES
sequence
1..164891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-444C12"
/clone_1b="RP11-11.2"
1..34236
/note="assembly-fragment:00027"
fragment_chain:1
clone_end:77
vector_side:left"
34337..47372
/note="assembly-fragment:00312"
fragment_chain:1"
47473..67025
/note="assembly-fragment:01923"
fragment_chain:1"
67126..71239
/note="assembly-fragment:02094"
fragment_chain:1"
71340..83903
/note="assembly-fragment:02213"
fragment_chain:1"
84004..91357
/note="assembly-fragment:01292"
fragment_chain:1"
91458..96452
/note="assembly-fragment:01391"
fragment_chain:1"
96553..148189
/note="assembly-fragment:02045"
fragment_chain:1"
148290..159778
/note="assembly-fragment:00427"
fragment_chain:1"
159879..164891
/note="assembly-fragment:00773"
fragment_chain:1"
clone_end:56
vector_side:right"
BASE COUNT 48361 a 32179 c 32357 g 51067 t 907 others
ORIGIN
Query Match 36.1%, Score 608; DB 2; Length 164891;
Best local similarity 100.0%; Pred. No. 1,2e-143;
Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1676 AGCTACAGATATATCATCTGATTAAGTATGATTCAGACAGAGTTCAA 1135
Db 49139 AGCTACAGATATATCATCTGATTAAGTATGATTCAGACAGAGTTCAA 49050
OY 1136 TGAATCTCTCAAGTAACTACTGCTGATCCCAAGGAGCACTGAGGAAGT 1195
Db 49019 TGAATCTCTCAAGTAACTACTGCTGATCCCAAGGAGCACTGAGGAAGT 48990
OY 1196 CTTTGTGTTAAACCAAGAACATTACTTTGAAATGGACAGATCTTTCAATTCAT 1255
Db 48939 CTTTGTGTTAAACCAAGAACATTACTTTGAAATGGACAGATCTTTCAATTCAT 48930
```


QY 1256 TCAGCGTGTGATAGGTGATCTGAAATCAGAAATATCCACATTCGACGAGTATCTTT 1315
DB 48929 TCAGCGTGTGATAGGTGATCTGAAATCAGAAATATCCACATTCGACGAGTATCTTT 48870
QY 1316 GTTATATCTCCACAGACTCCGCGACAGACACCTACTCTGATGAAGCTCTGCTCTTG 1375
DB 48869 GTTATATCTCCACAGACTCCGCGACAGACACCTACTCTGATGAAGCTCTGCTCTTG 48810
QY 1376 TCCATATATTCATATCAGACAGACATTCCTGCGATTCACATTTTAAATATATGGA 1435
DB 48809 TCCATATATTCATATCAGACAGACATTCCTGCGATTCACATTTTAAATATATGGA 48750
QY 1436 GTGAGATAGAGAGACGACGCTGCTCAATACCTAGCGCTGATTTTGTGAGATAA 1495
DB 48749 GTGAGATAGAGAGACGACGCTGCTCAATACCTAGCGCTGATTTTGTGAGATAA 48690
QY 1496 ATAAATCATTCATCCCTTTTGTGATTAATAATTTCTAAATGATTTTACATTCCT 1555
DB 48689 ATAAATCATTCATCCCTTTTGTGATTAATAATTTCTAAATGATTTTACATTCCT 48630
QY 1556 GTAGGGGGGATTAATAATGATATAGTATTAATTAATAATGATTTTCTGAGGG 1615
DB 48629 GTAGGGGGGATTAATAATGATATAGTATTAATTAATAATGATTTTCTGAGGG 48570
QY 1616 GCGATATTAATAATGATTTTAGACTTCTGTAAGGGGGGATTAATAATGATTTTCTGAGGG 1675
DB 48569 GCGATATTAATAATGATTTTAGACTTCTGTAAGGGGGGATTAATAATGATTTTCTGAGGG 48510
QY 1676 ACTGGGTA 1683
DB 48509 ACTGGGTA 48502

RESULT 14
AF039401
LOCUS AF039401 35278 bp DNA PRI 14-DEC-1998
DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene,
complete cds.
ACCESSION AF039401
VERSION AF039401.1 GI:4009459
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 35278)
AUTHORS Gruber,A.D., Elble,R.C., Ji,H.L., Schneur,K.D., Fuller,C.M. and
Pauli,B.U.
TITLE Genomic cloning, molecular characterization, and functional
analysis of human CLCA1, the first human member of the family of
Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 35278)
AUTHORS Gruber,A.D., Elble,R. and Pauli,B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Department of Pathology, College of
Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

FEATURES
Source
1..35278
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
1..1617
/gene="hCLCA1"
promoter
1..33522
/gene="hCLCA1"
gene
1..33522
/gene="hCLCA1"
repeat_region
767..1265
/rpt_family="L1 transposable element"
TATA_signal
1587..1596
/gene="hCLCA1"

/note="putative"
1618..1708
/gene="hCLCA1"
/number=1
join(1618..1708,1962..2383,6667..6807,6979..7126,6966..9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33522)
/gene="hCLCA1"
/product="calcium-dependent chloride channel-1"
1962..2383
/gene="hCLCA1"
/number=2
join(2222..2383,6667..6807,6979..7126,6966..9801,15454..15631,18591..18809,19783..20010,22253..24522..24628,26644..26859,27447..27708,28766..31840..32079,32919..33510)
/gene="hCLCA1"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95429.1"
/db_xref="GI:4009460"
/translation="MGPRSSVFLILHLLEGLNSLIQLNNNGEIVAVIDNPV
EDETLIQIKDWVQASLYLFEPATGRKRFKRVNAILIPETWTKADYVRKLETYKNA
DVLVAESTPEPDDEPTBQMGNCCKGKGEIHLTPDLAKGLAEYGPQKAFVHMAL
LHWGVFDEYNDEKFEYLSNGRIQAVKSGSIGTGVVVKCGGSCYTRKCFENYGL
YKGGVEFYLOSROTERKASIMFAOHVDSIYEFCTEONHNKEAPNKNOKNCTSTWEVI
RSDPEFKTTPMTQIPNPPTSLIQIGRIVGLVLDKSGSNATGGRNRLNAGOLF
LQTVELGSRVGVATIDSMANVOSLIQINSQSDRPTLAKRLPAASGGTSITSGLSA
FTVIRKRPDSEIVLITDGEDNTISGFENVKOSGAIIRFVALGPSAOLLEELSK
MTGGLQVADSDOVNNGLLDAPGALSNGAVSOSIQLSEGLQLONSQWNGVIV
DSTVGKDFLTLTWTQPPQILLMPDSGKOGFVVDKMTKAYVQIIPGIAVNGTWY
SQASSQTLITVTSRASNAFLPITVTSKTKDTSKPSPLVYVANIISQVAPLIRA
SVTALIESVNGKVTLELDNGAGADATYDQGVSRFETPTDNRGSRVAVLGGVN
AARRVVIOQSALYIRGVIENDELQWNPPEIRKNDQVQHOVCFESRSSGSEYAS
DVPNAPIDLPFGQITDLNATIHGSLINLWTAPGDDYDGHAKITIRISTSLD
LNDKRNESIQVWTFALIPKANSSEVFLKPNITFENCTDLFLAIQAVDKVLSSEI
SNIRVSLFIPQTPPEPSPDSTAPCPNINHINSTIRPIHILKIMKWIQELQSLIA
"
6667..6807
/gene="hCLCA1"
/number=3
6979..7126
/gene="hCLCA1"
/number=4
9696..9801
/gene="hCLCA1"
/number=5
15454..15631
/gene="hCLCA1"
/number=6
18591..18809
/gene="hCLCA1"
/number=7
19783..20010
/gene="hCLCA1"
/number=8
22253..22427
/gene="hCLCA1"
/number=9
24522..24628
/gene="hCLCA1"
/number=10
26644..26859
/gene="hCLCA1"
/number=11
27447..27708
/gene="hCLCA1"
/number=12
28766..28936
/gene="hCLCA1"
/number=13
31840..32079

Oy	412	CTGGAAATACAGCT-----GCAGGCAAGCTCACAAACCTTGACCCGTGATGCA	462
Db	1718	CTTGGACTTACAGCTCTTAATAATATCCACAGCTCTCAAAATCTAACAGGACATGA	1777
Oy	463	CGTCCCGTGGTCCAAATGCTACCCCTGCTCCAAATTCAGTGCATTCGAAAGCAAGG	522
Db	1778	CCACTCCAGCAAGAAGTCTACTACACCCCCAGTACTGCACACGCTCAGATGATCAA	1837
Oy	523	ACACCAACAATTTCCCGAGCCCTCTGGTAGTTTATGCAAAATTTGCGCAAGAGCCCTCC	582
Db	1838	ATACAGACACTACCCCTAGGCCAGTGATGTTTATGCACAGATCAGTCAGTCAAGGTTTTGC	1897
Oy	583	CAATTCCTCAGGCGCAGTGTACAGACCCCTGATTTGAATCACTGAAATGAAAAACATTTACT	642
Db	1898	CTGTACTGGGAATCAATGATCACTGATATGAAACTGAAATGCGCATCAAGTAACAC	1957
Oy	643	TGGAACTACTGGATTAATGGAGAGGTGATGCTCTAAAGATGACGGTGTCTCTCA	702
Db	1958	TGGAGCTCTGGGACAAATGGTGGCAGGTGCTGATATCTCAAGATGATGGCATCTATTTCA	2017
Oy	703	GGTATTTACAACTTATGACAGCAATGATAGATACAGTGTAAAGATGGGGCTCTGGAG	762
Db	2018	GATATTTACAGATTTACCTGTGAATATGGTATGATACAGTTTAAAGTACATGCCGAGCA	2077
Oy	763	GAGTTAACGACGACGACGAGAGATATACCCAGAGAGTGAGACTGTATCATACTCG	822
Db	2078	GAACAACAGGGCTAGGCTAAGTTTAAAGCAACACAGCAACCCCTGTATATACAG	2137
Oy	823	GCTGATTTGGAATGATGAATCAATGCAATGCAACAGACCTGAAATTAATGAGATG	882
Db	2138	GCTTATTTGAAATGTATTAATATATCTGACACCCAGCCTGAGTCAAGATGAC	2197
Oy	883	ATGTTCAACACAAACCAAGATGTGTTTCAGAGAAACATCTCGGAGGCTCATTTGTGCTT	942
Db	2198	TGGCAAAAGCTGAATATGAAAGACTTTAGCAGACTAACCTCTGAGGGTCAATTACTGTAT	2257
Oy	943	CTGATGTCCTCCAAATGCTCCCATACCTGATGTCCTTCCCACTGCGCAATCACCACTGTA	1002
Db	2258	CAGAGGCTCCTCCGTGTAATCACCTCTGTGCTCCACCACTAAATTAATTAAGCACTTG	2317
Oy	1003	AGGGGGAATTCACGSGGGGAGTCTCATTAATCTGACTTGGACAGCTCCTGGGATGAT	1062
Db	2318	AAGCAAAATTTTAAAGAAG---ATCACATTCACTTTCATGACAGAGCCCTGCGAATGTCC	2374
Oy	1063	ATGACATGGAACAGCTCACMAATATATCATTTGGAATAGTACAAGTATCTTGTATCA	1122
Db	2375	TAGATTAAGGAAAAAGCCACAGCTACATTAATAAGATTAAGTAAGTTCCTCGATCTCC	2434
Oy	1123	GAGACAAGTTCATGAATCTCTTCAAGTGAATATCTACTGCTCTCATCCCAAGAGCA	1182
Db	2435	AAAAAAGTTTGGACAAGTCACTTGTAGTATCTTCCAGTCTTAAACCTTAAGGAGGCG	2494
Oy	1183	ACTGTGAGGAAGCTCTTTTCTTAAACACGAAAAACATCTTGTGAATGGCAGACATC	1242
Db	2495	GCTCAGATGAAATTTTGAATTTAAACCGAAGCCTTTTGAATAGAAATGGCACCACT	2554
Oy	1243	TTTTCATTTGCTATTCAGGCTGTGATTAAGGTGATCTGAAATCAAGAAATATCCAACTG	1302
Db	2555	TCTATATTTGCGAGTCCAAAGCATCAATGAAGCAATCTCAGAGGTTCTAATCATCG	2614
Oy	1303	CACGAGTATCTTTGTTATTTCTT	1325
Db	2615	CACAAGCAATCAAGTTTATTTCTT	2637

Search completed: April 4, 2002, 03:16:12
Job time: 61254 sec

ms Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:48 ; Search time 16681 Seconds
(without alignments)
1812.111 Million cell updates/sec

Title: US-09-049-696-18

Perfect score: 2813
Sequence: 1 GAATCAGCAGGAGATCTAC.....AAATTAATCATTCCTTA 2813

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estfun:
2: em_esthum:
3: em_estin:
4: em_estom:
5: em_estpl:
6: em_estcda:
7: em_estro:
8: em_estov:
9: em_hic:
10: qb_est1:
11: qb_est2:
12: qb_hic:
13: qb_gss:
14: em_gss_fun:
15: em_gss_hum:
16: em_gss_inv:
17: em_gss_pln:
18: em_gss_pro:
19: em_gss_rtd:
20: em_gss_vtl:
21: em_gss_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1739.8	61.8	2933	12 AK008659	AK008659 Mus muscu
2	1725.6	61.3	2915	12 AK007466	AK007466 Mus muscu
3	646.4	23.0	734	10 A1660957	A1660957 wf20D08.x
4	612.6	21.8	653	10 AM361534	AM361534 QV2-CT026
5	567.8	20.2	598	11 BE927204	BE927204 RCL-CT024
6	565	20.1	633	10 AM753451	AM753451 QV2-CT026
7	557	19.8	625	10 AM361532	AM361532 QV2-CT026
8	553	19.7	657	10 AM361521	AM361521 QV2-CT026
9	551.4	19.6	621	10 AM361520	AM361520 QV2-CT026
10	535.2	19.0	561	10 A1802756	A1802756 wf18A04.x
11	521.8	18.5	712	10 AM009763	AM009763 ws87D04.x
12	520	18.5	652	10 AM753449	AM753449 QV2-CT026

13	516.8	18.4	528	11 BF083270	BF083270 RCL-CT024
14	487.2	17.3	561	10 AM361523	AM361523 QV2-CT026
15	487	17.3	504	10 A1802693	A1802693 wf17C01.x
16	481.4	17.1	716	10 A1660234	A1660234 wf68G02.x
17	473.6	16.8	480	10 AA581198	AA581198 nd38C07.x
18	471.2	16.8	501	10 A1721275	A1721275 as82H08.x
19	461.8	16.4	654	10 AM361522	AM361522 QV2-CT026
20	444	15.8	460	10 AM750670	AM750670 CM2-CN03
21	439.4	15.6	806	11 BC962080	BC962080 602826824
22	434	15.4	455	11 BF083321	BF083321 RCL-CT024
23	432.4	15.4	496	10 AM753448	AM753448 QV2-CT026
24	421.8	15.0	661	11 BF581041	BF581041 602100173
25	421.8	15.0	935	11 BF578919	BF578919 602095492
26	421.2	15.0	466	11 BF195179	BF195179 RST14358
27	411.8	14.6	906	11 BF581342	BF581342 602100726
28	399	14.2	767	11 BF578246	BF578246 6020931673
29	397.6	14.1	704	11 BC965023	BC965023 602831668
30	391.6	13.9	745	11 BC963539	BC963539 602831217
31	384.6	13.7	890	11 BF578833	BF578833 602094876
32	379.4	13.5	911	11 BF578693	BF578693 602093175
33	378.8	13.5	876	11 BF234977	BF234977 602027529
34	369.4	13.1	974	11 BC962464	BC962464 602827322
35	369.2	13.1	998	11 BF583412	BF583412 602097421
36	358.8	12.8	435	10 AM844161	AM844161 RCL-CT004
37	352.8	12.5	356	10 AM361528	AM361528 QV2-CT026
38	349.8	12.4	747	11 BF580957	BF580957 602100659
39	347.8	12.4	378	10 AA296955	AA296955 EST112726
40	342.8	12.2	812	11 BC968878	BC968878 602836186
41	336.8	12.0	646	10 AA691335	AA691335 vs14h06.x
42	326.4	11.6	813	11 BG216984	BG216984 RST36685
43	326	11.6	479	10 AA541829	AA541829 vJ01F01.x
44	323.8	11.5	469	10 BE136935	BE136935 ug57e12.x
45	322.2	11.5	335	10 A1721121	A1721121 as73d08.x

ALIGNMENTS

RESULT 1
LOCUS AK008659 2933 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210006G11, full insert sequence.
ACCESSION AK008659
VERSION AK008659.1 GI:12842987
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library
clone:2210006G11.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2933)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Methods in enzymology. 303, 19-44 (1999)
JOURNAL JOURNAL
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (bases 1 to 2933)
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome research. 10 (10), 1617-1630 (2000)
JOURNAL JOURNAL
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (bases 1 to 2933)
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

D	b	809	CCAATGACCCAAACCAACGATGCAATCTCCGAAGACAGTGGGAAGTATATCCAGGAATCT	866
O	y	862	GAGACCTTTAAGAAAACCACTCCTATGACAAACAGACCCACCAAAATCCACCTTTCATTTG	921
D	b	869	GAGACCTTCAAGCAAAACCACTCCCATGACAGCCACCCACCTGACACCCACTTTCACAG	928
O	y	922	CTGAGAGTTGGACCAAAATATGTGTGTTATATGTCCTTGAACAAATCTGGAAGCATGGGCAGT	981
D	b	929	CTGCAAAATTTGGACAAACAAATTTGTGTCTTAAGTCTTGTATTAAGTCCGGAGCATGCTGAAC	988
O	y	982	GGTAAACCGCCCAATTCACATGAAATCAAGACAGCCAGCACCTTTCTCGTGGAGACACTTGGAG	1041
D	b	989	GATGATGCTCTTAACCCAAATGAATCAGGCGAACGCCGCTTTCTGCTGACAGCTGTGGAG	1048
O	y	1042	CTGGGGTCTCGGTTTGGGATGTTGACATTTTACAGTGTGCCCATGTACAAAGTGAACATC	1101
D	b	1049	CAGGATCTCTCGTGGGATGGTACCTTTGACAGTGTGCCATATTTACAAAGCCAGATC	1108
O	y	1102	ATACAGATTAACCACTGGCACTGACAGGGACACACTCGCCAAAAGATTACCTCAGCAGCT	1166
D	b	1109	AAACAGTTAAACATGTGTGCTGACAGAGATCTCTATTAAGCACTTACCACAGATCT	1166
O	y	1162	TCAGAGGAGAGCTCATCTGACAGCGGACCTTCGATCGGACATTTACGTGATTTAGAGAA	1221
D	b	1169	GCAGAGGAGGACATTTATATGCTCTGGCTTTCGACACAGCATTTACAGTGTAAAGAGAG	1222
O	y	1222	TATCCAACTGATGATCTGAAATTTGTGCTGTCAGCGATGGGGAAGACAACTATAAT	1281
D	b	1229	TATCCAACTGATGATCTGAAATTTGTGCTGTCAGCGATGGGGAAGACAACTATAAT	1288
O	y	1282	GGGTGCTTTAACGAGTCAACAAAGTGTGCTCATATCCACACAGTGGCTTTGGGGCC	1341
D	b	1289	AGCTGCTTTGACCGTGTGAAGCAGAGGGGGCCATCATCACTGACGTGGGACCG	1348
O	y	1342	TCCTGAGCTCAAGAACTAGAGAGAGCTGTCCAAAATACAGAGGCTTACAGCATATGCT	1401
D	b	1349	GCTGCCGCTTAAGAGCTTGTAGCAAGCTGTCCAAAATACAGAGGCTGTGACAGCATCTCT	1408
O	y	1402	TCGATCAAGTTTCGAAACAATTGGCTCATTTGATGCTTTTGGGGCCCTTTCATCAGGAAT	1461
D	b	1409	TCGGATCAGTTTCAGAACAAATGGCTTGTGATGTTCGATTCGACAGCACTCTCTCAGGAAT	1466
O	y	1462	GGACATCTCTCTCAGCGCTCCATTCACAGTTGAAAGTAAAGGATTTAACCTCCAGAACAC	1521
D	b	1469	GGCGGATCGCTGACACTCTCATTCACAGCTGGAAGAGGGAGTATATCTCCAGAAATAC	1528
O	y	1522	CAGTGAATGAATGACAGAGTATGCTGGAGCAGCAGCCGTGGGAAAGACACTTTGTTCTT	1581
D	b	1529	CATATGATGAATGGCTATGATGATCTGATGACAGCTCGGTGGGCAAGACACTTGTCTT	1588
O	y	1582	ATCACCTGACAAACGACAGCCTCCCAAAATCTCTCTGGGATCCCAAGTGGACAGAACAA	1641
D	b	1589	ATCACCTGACAAACGACAGCCTCCCAAAATCTCTCTGGGATCCCAAGTGGACAGAACAA	1648
O	y	1642	GGTGGCTTTGATGAGACAAAACACCAAAATGGCTTACTCCAAATCCAGGACTTGGT	1701
D	b	1649	AATGGTTTATATCTAGACACAACACACTAAGGTGGCTTACTCCAAATCCAGGACAGCT	1708
O	y	1702	AAGGTGGCACTTGGAAATACAGTCTGACAGAACGTCACAAACCTTGAACCTGAGCTGTC	1766
D	b	1709	AAGGTGGCTTTGGAAATACAGATTCAAAGGAGCTCACAGACTCTCACTTGTAGTGC	1766
O	y	1762	ACGTCCGTGGCTCAATTCGTACCTGCTCCCAATTTACAGTGAAGTCCAAAACGAACAG	1821
D	b	1769	ACGTCCGTGGAGCAAGTGTGACACTGCTCTCTATTAACAGTGAACCCGGTACTGAATTAAG	1822
O	y	1822	GACACCAAGCAATTTCCCAAGCCTCTGTGATGTTATGCAAAATTTTCGCCAAGAGCTTC	1881
D	b	1829	AACACAGGAAATTTCCCAAGCCTCTGTAACTGATGCAAGCACTTTCGCCAAGAGACTTCG	1888
O	y	1882	CCATTTCCAGGGCACTGTGCAAGCCGTGATGAATCAAGTGAATGGAATGAACACTTACC	1941
D	b	1889	CTATTTCTCAGGGCACTGTGCAAGCCGTGATGAATCAAGTGAATGGAATGAACACTTACC	1948

[illegible]

AK007466.1	2915 bp	mRNA	HTC	05-JUL-2001
Mus musculus	10 day old male pancreas	cDNA	RIKEN full-length	
enriched library,	clone:1810012P03,	full insert	sequence.	
AK007466				
AK007466.1	GI:12841032			
Cap trapper.				
Mus musculus (strain:C57BL/6J)	10 day old male pancreas	cDNA	to	
mRNA, clone:lib:RIKEN full-length	enriched mouse	cDNA library	clone:1810012P03.	

LOCUS AM753451 633 bp mRNA EST 28-APR-2000
DEFINITION QV2-CT0261-261099-011-d05 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM753451
VERSION AM753451.1 GI:7668383
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 633)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R.,
Nagai,M.A., da Silva,W.J.F., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
2020263
JOURNAL Contact: Simpson A.J.G.
MEDLINE Laboratory of Cancer Genetics
COMMENT Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l-QV2&t2-QV2-CT0261-
261099-011-d05&t3-1999-10-26&t4-1)
Seq primer: puc 18 forward
High quality sequence start: 66
High quality sequence stop: 90.
Location/Qualifiers
1. 633
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
, 716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 136 a 151 c 161 g 185 t
ORIGIN
Query Match 20.1%; Score 565; DB 10; Length 633;
Best Local Similarity 95.9%; Pred. No. 2,6e-130;
Matches 580; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1495 AGTAAGGATTACCTCCAGACAGCCAGTGAATGACAGTGTGTGACAGC 1554
DB 612 AGATAAGATTACCTTACCTAGACAGCCAGTGAATGACAGTGTGTGACAGC 553
QY 1555 ACCGTGGGAAGGACACTTTTCTTATCACTGAGACACGCCACCTCCCAATCTT 1614
DB 552 ACCGTGGGAAGGACACTTTTCTTATCACTGAGACACGCCACCTCCCAATCTT 493
QY 1615 CTCTGGGATCCAGTGGACAGACAGGAGTGTGAGTGACAAACACCAAAATG 1674
DB 492 CTCTGGGATCCAGTGGACAGACAGGAGTGTGAGTGACAAACACCAAAATG 433
QY 1675 GCCTACTCCCAATCCAGGAGTGTGAGTGACCTTGAATATCACTGTCAAGCA 1734
DB 432 GCCTACTCCCAATCCAGGAGTGTGAGTGACCTTGAATATCACTGTCAAGCA 373
QY 1735 AGCTCAACAACTTGACCTGACTGTACGTCGCCGTGCTCAATGCTACCTGCTTCCA 1794

DB 312 AGCTCAACAACTTGACCTGACTGTACGTCGCCGTGCTCAATGCTACCTGCTTCCA 313
QY 1735 ATTACAGTACTCCCAATCCAGGAGTGTGAGTGACCTTGAATATCACTGTCAAGCA 1854
DB 312 ATTACAGTACTCCCAATCCAGGAGTGTGAGTGACCTTGAATATCACTGTCAAGCA 253
QY 1855 TATGCAAAATATTCGCCAGAGAGCCCTCCCAATTCACAGGCGAGTGTACAGCCCTGATT 1914
DB 252 TATGCAAAATATTCGCCAGAGAGCCCTCCCAATTCACAGGCGAGTGTACAGCCCTGATT 193
QY 1915 GAATCAGTGAATGGAAAACAGTACTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAT 1974
DB 192 GAATCAGTGAATGGAAAACAGTACTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAT 133
QY 1975 GCTACTAGATGAGCGGTGTCTACTACAGTATTCACAACTTATGACAGCAATGTTGA 2034
DB 132 GCTACTAGATGAGCGGTGTCTACTACAGTATTCACAACTTATGACAGCAATGTTGA 73
QY 2035 TACAGTGTAAAAGTGGCGGCTCTGGAGAGTAAAGCAGCCAGAGAGATGATACC 2094
DB 72 TACAGTGTAAAAGTGGCGGCTCTGGAGAGTAAAGCAGCCAGAGAGATGATACC 13
QY 2095 CAGCA 2099
DB 12 CAGCA 8
RESULT 7
AM361532
LOCUS AM361532 625 bp mRNA EST 04-FEB-2000
DEFINITION QV2-CT0261-261099-011-e03 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361532
VERSION AM361532.1 GI:6866286
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 625)
HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l-QV2&t2-QV2-CT0261-
261099-011-e03&t3-1999-10-26&t4-1)
Seq primer: puc 18 forward
High quality sequence start: 93
High quality sequence stop: 608.
Location/Qualifiers
1. 625
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
, 716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 172 a 154 c 150 g 149 t
ORIGIN


```

Query Match      19.8%;   Score 557;   DB 10;   Length 625;
Best Local Similarity 99.1%;   Pred. No. 2.6e-128;
Matches 560; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Oy	1235	GATCTGAAATTTGTCCTCTGACGGATGGGGAACAAACACTTAATAGTGGCTGCTTTAAAG	129
Db	61	GATCTGTAATTTGTCTCTCTGACGGATGGGGAACAAACACTTAATAGTGGCTGCTTTAAAG	120
Oy	1295	AGGTCAAAACAAAGTGTGTCATCATCCACACAGATCGCTTTGGGGCCCTTGGAGCTCAAG	135
Db	121	AGGTCAAAACAAAGTGTGTCATCATCCACACAGATCGCTTTGGGGCCCTTGGAGCTCAAG	180
Oy	1335	AACTAGAGGAGCTGTCCAAAATACAGAGAGTTTACAGACATATGCTTCAGATCAAGTTC	1414
Db	181	AACTAGAGGAGCTGTCCAAAATACAGAGAGTTTACAGACATATGCTTCAGATCAAGTTC	240
Oy	1415	AGAACAAATGGCCATATGATGATGCTTTTGGGGCCCTTTCATACAGAAATGAGAGCTGTCTTC	1478
Db	241	AGAACAAATGGCCATATGATGATGCTTTTGGGGCCCTTTCATACAGAAATGAGAGCTGTCTTC	300
Oy	1475	AGGCTCTCATCAGCTTGAGAGATTAAGGGAATTAACCTCCAGAAACAGCCAGTGGATGAATG	1534
Db	301	AGGCTCTCATCAGCTTGAGAGATTAAGGGAATTAACCTCCAGAAACAGCCAGTGGATGAATG	360
Oy	1535	GCACAGTGAATCTGTGACAGCAGCGTGGGAAAGGACACTTTGTTCTTATCACCTGGACAA	1594
Db	361	GCACAGTGAATCTGTGACAGCAGCGTGGGAAAGGACACTTTGTTCTTATCACCTGGACAA	420
Oy	1595	CGCAGCCTCCCAAATTCCTTCTGTGGGATCCCACTGTGGAAGCAAGAGTGGCTTTGTAG	1655
Db	421	CGCAGCCTCCCAAATTCCTTCTGTGGGATCCCACTGTGGAAGCAAGAGTGGCTTTGTAG	480
Oy	1655	TGAGCAAAAAACACCAAAATGGCTACCTCCAAATCCAGGACATTTGCTAAGTTGGCACATT	1714
Db	481	TGAGCAAAAAACACCAAAATGGCTACCTCCAAATCCAGGACATTTGCTAAGTTGGCACATT	540
Oy	1715	GGAAATTCAGCTGTGCAGAGCAAGCTCAAAACCTTGGACCTGACTGACAGTCCCGTGGCT	1714
Db	541	GGAAATTCAGCTGTGCAGAGCAAGCTCAAAACCTTGGACCTGACTGACAGTCCCGTGGCT	600
Oy	1775	CCAAATGCTACCTCTCCCTCCAAATTAC	1799
Db	601	CCAAATGCTACCTCTCTCCAAATTAC	625

RESULT	8
AM361521	
LOCUS	AM361521 657 bp mRNA
DEFINITION	QVJ-C70261-261099-011-f03 C70261 Homo sapiens EST
ACCESSION	AM361521
VERSION	AM361521.1 GI:6866275
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	1 (bases 1 to 657)
TITLE	HCGP http://www.ludwig.org.br/ORESTES .
JOURNAL	The FAPESP/LICR Human Cancer Genome Project
COMMENT	Unpublished (1999)
	Contact: Simpson A.J.G.

```
261099-011.f038ct3-1999-10-26ct4=1
Seq primer: puc 18 forward
High quality sequence start: 123
High quality sequence stop: 656.
Location/Qualifiers
FEATURES
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT      175 a      160 c      158 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
```

QY	1138	GCATTTCTCTGATTTTGGGAAGAAATTCACAACGATGAG-TCGTGAAATGTCCTCTAC	1256
Db	1139 <td>GCATTAATGTCGATTTGGGAACCAATTAATCTAACGATGAGATTCGTAAATGTCCTCTAC</td> <td>128</td>	GCATTAATGTCGATTTGGGAACCAATTAATCTAACGATGAGATTCGTAAATGTCCTCTAC	128
QY	1137 <td>GGATGGGGAAGACACACTTAATAGGCGTCCTTAAACGAGTCCAAACAAAGTGTCGAT</td> <td>1316</td>	GGATGGGGAAGACACACTTAATAGGCGTCCTTAAACGAGTCCAAACAAAGTGTCGAT	1316
Db	1129 <td>GGATGGGGAAGACACACTTAATAGGCGTCCTTAAACGAGTCCAAACAAAGTGTCGCT</td> <td>188</td>	GGATGGGGAAGACACACTTAATAGGCGTCCTTAAACGAGTCCAAACAAAGTGTCGCT	188
QY	1117 <td>CATCCACACAGTGGCTTGGGGGCCCTCGAGTCGCAAGCACTAGAGGAGCTGTCCAAAT</td> <td>1376</td>	CATCCACACAGTGGCTTGGGGGCCCTCGAGTCGCAAGCACTAGAGGAGCTGTCCAAAT	1376
Db	1109 <td>CATCCACACAGTGGCTTGGGGGCCCTCGAGTCGCAAGCACTAGAGGAGCTGTCCAAAT</td> <td>248</td>	CATCCACACAGTGGCTTGGGGGCCCTCGAGTCGCAAGCACTAGAGGAGCTGTCCAAAT	248
QY	1177 <td>GACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCCAGAACAAATGGCTCATTTGATGC</td> <td>1436</td>	GACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCCAGAACAAATGGCTCATTTGATGC	1436
Db	1249 <td>GACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCCAGAACAAATGGCTCATTTGATGC</td> <td>308</td>	GACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCCAGAACAAATGGCTCATTTGATGC	308
QY	1437 <td>TTTTGGGCCCCCTTTCATCAGGAATGAGCGTGTCTCTCAGCCCTCCATCCAGCTTGAGAG</td> <td>1496</td>	TTTTGGGCCCCCTTTCATCAGGAATGAGCGTGTCTCTCAGCCCTCCATCCAGCTTGAGAG	1496
Db	1309 <td>TTTTGGGCCCCCTTTCATCAGGAATGAGCGTGTCTCTCAGCCCTCCATCCAGCTTGAGAG</td> <td>368</td>	TTTTGGGCCCCCTTTCATCAGGAATGAGCGTGTCTCTCAGCCCTCCATCCAGCTTGAGAG	368
QY	1497 <td>TAAAGGATTAACTCCCTCAGAACAGCCAGTGGATGAAATGGCACAGTATGATGGGACAGAC</td> <td>1556</td>	TAAAGGATTAACTCCCTCAGAACAGCCAGTGGATGAAATGGCACAGTATGATGGGACAGAC	1556
Db	1369 <td>TAAAGGATTAACTCCCTCAGAACAGCCAGTGGATGAAATGGCACAGTATGATGGGACAGAC</td> <td>428</td>	TAAAGGATTAACTCCCTCAGAACAGCCAGTGGATGAAATGGCACAGTATGATGGGACAGAC	428
QY	1527 <td>CGTGGGAAAGACACTTTGTGTTTATTCACCTGGACACAGCAGCCCTCCCAATCTTCT</td> <td>1616</td>	CGTGGGAAAGACACTTTGTGTTTATTCACCTGGACACAGCAGCCCTCCCAATCTTCT	1616
Db	1409 <td>CGTGGGAAAGACACTTTGTGTTTATTCACCTGGACACAGCAGCCCTCCCAATCTTCT</td> <td>488</td>	CGTGGGAAAGACACTTTGTGTTTATTCACCTGGACACAGCAGCCCTCCCAATCTTCT	488
QY	1617 <td>CTGGGATCCCACTGGACAGAGCAAGCAGTGCGCTTGTAGTGACCAAAAACACCAAAATGCG</td> <td>1676</td>	CTGGGATCCCACTGGACAGAGCAAGCAGTGCGCTTGTAGTGACCAAAAACACCAAAATGCG	1676
Db	1409 <td>CTGGGATCCCACTGGACAGAGCAAGCAGTGCGCTTGTAGTGACCAAAAACACCAAAATGCG</td> <td>548</td>	CTGGGATCCCACTGGACAGAGCAAGCAGTGCGCTTGTAGTGACCAAAAACACCAAAATGCG	548
QY	1677 <td>CTACCTCCCAATCCCGAGGATGCTTAAGTGTGGACCTTGGAAATTAACAGTCTCAAGCAAG</td> <td>1736</td>	CTACCTCCCAATCCCGAGGATGCTTAAGTGTGGACCTTGGAAATTAACAGTCTCAAGCAAG	1736
Db	1549 <td>CTACCTCCCAATCCCGAGGATGCTTAAGTGTGGACCTTGGAAATTAACAGTCTCAAGCAAG</td> <td>608</td>	CTACCTCCCAATCCCGAGGATGCTTAAGTGTGGACCTTGGAAATTAACAGTCTCAAGCAAG	608
QY	1727 <td>CTCACAACCTTGACCCCTGACTGTCACTGTCGCGGTGCGTCAATGCTAC</td> <td>1785</td>	CTCACAACCTTGACCCCTGACTGTCACTGTCGCGGTGCGTCAATGCTAC	1785
Db	609 <td>CTCACAACCTTGACCTGACTGTCACTGTCGCGGTGCGTCAATGCTAC</td> <td>657</td>	CTCACAACCTTGACCTGACTGTCACTGTCGCGGTGCGTCAATGCTAC	657

RESULT 9

AM361520

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

EST

621 bp

mrna

QV2-CT0261-261099-011-e07

CT0261

Homo sapiens

CDNA, mRNA sequence

GI:6866274

EST

04-FEB-2000

Source	Organism	Human
REFERENCE	Human	Human sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiina; Homidae; Homo.	
JOURNAL	1 (bases 1 to 621)	
COMMENT	HCGE http://www.ludwig.org.br/ORESTES . The FAPESP/LICR Human Cancer Genome Project Unpublished (1999) Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?el=QV2&t2=QV2-CT0261 261099-011-e07&t3=1999-10-26&t4=1) Seq primer: puc 18 forward High quality sequence start: 88 High quality sequence stop: 613. Location/Qualifiers 1. 621 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="CT0261" /dev_stage="Adult" /note="Organ: colon; Vector: puc18; Site.1: Sma1; Site.2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions." BASE COUNT 166 a 151 c 150 g 154 t ORIGIN	
Query Match	19.6%; Score 551.4; DB 10; Length 621;	
Best Local Similarity	96.2%; Pred. No. 6.6e-127;	
Matches	586; Conservative 0; Mismatches 21; Indels 2; Gaps 2	
OY	1187 GGCCTTCGATCGGCATTTACTGTGATTTAGGACAATAATCCAACTGATGCTGAATTC	1246
Db	11 GGATGTGATCTGTCATTTACTGTGATTTAGGACAAGATTATTAACATGATGCTGTAATTC	69
OY	1247 TGCCTCTGACGATGGGGAAGACAACACATATTAAGTGGGCTTTAAACGAGGTCAAACAA	1306
Db	70 TGCCTCTGACGATGGGGAAGACAACACATATTAAGTGGGCTTTAAACGAGGTCAAACATA	129
OY	1307 GTGTGTCATCATTCCACACAGTCGGCTTTGGGGCCCTCTGACCTCAAGAACTTAGAG-GAG	1365
Db	130 GTGTGTCCTCTCATTCCACACAGTCGGCTTTGGGGCCCTCTGACCTCAAGAACTTAGAGTGA	189
OY	1366 CTGTCCAAATGACAGGAGGTTTACAGACATATGCTTACAGATCAAGTTTCAGAACTATGGC	1425
Db	190 CTGTCCAAATGACAGGAGGTTTACAGACATATGCTTACAGATCAAGTTTCAGAACTATGGC	249
OY	1426 CTCATTGATGCTTTTGGGGCCCTTTCATCAGGAATAATGAGAGTGTCTCAGCGCTTCATC	1485
Db	250 CTCATTGATGCTTTTGGGGCCCTTTCATCAGGAATAATGAGAGTGTCTCAGCGCTTCATC	309
OY	1486 CAGCTTTGACAGTAAGGATTTAACTCTCCAGAAACAGCCAGTGTGATGATGACAGTATC	1545
Db	310 CAGCTTTGACAGTAAGGATTTAACTCTCCAGAAACAGCCAGTGTGATGATGACAGTATC	369
OY	1546 GTGACACAGACCGTGGGAAAGACACTTTGTTTTCATCACCTGGAGCAACAGAGCTCCG	1605
Db	370 GTGACACAGACCGTGGGAAAGACACTTTGTTTTCATCACCTGGAGCAATGACAGCTCCG	429
OY	1606 CAAATCTCTCTCTGGGATCCAGTGGACAGAAAGCAAGGTGGCTTTGTATGTGACAAAAAC	1665

	DB	489	CANAATCTTCTGTGGATCCGATCCGACAGCAAGCAGGCGCTTTGTATGGACAATAAAC	489
	QY	1666	ACCAAATGGCTTACCTCCAAATCCGAGCATTTGCTAAGTTGGACTTGGAATACACT	1725
	Db	489	ACCAAATGGCTTACCTCCAAATCCGAGCATTTGCTAAGTTGGACTTGGAATACACT	549
	QY	1726	GTGCAGCAGCAGTCACAAACCTTGACCTGTGATGCTGCCGTCGCTCAATGCTACC	1785
	D5	590	GTGCAGCAGCAGTCACAAACCTTGACCTGTGATGCTGCCGTCGCTCAATGCTACC	609
	QY	1786	GTGCCTCCA 1794	
	Db	610	TGGCTCCA 618	
	RESULT	10		
	LOCUS	A1802756/c		
	DEFINITION	A1802756	561 bp	mRNA
	VERSION	wf18a04.x1 Soares_Dieckgraefe_colon_NHUC Homo sapiens cDNA clone	EST	18-DEC-1999
	KEYWORDS	IMAGE:2350926 3' similar to TR:088826 088826 GOB-5 PROTEIN. ; , mRNA sequence.		
	SOURCE	A1802756.1 GI:5368228		
	ORGANISM	human.		
	REFERENCE	Homo sapiens		
	AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
	COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index unpublished (1997)		
	FEATURES	Contact: Robert Strausberg, Ph.D. Email: rgs@biomail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert Length: 840 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 398. Location/Qualifiers 1..561 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2350926" /clone_lib="Soares_Dieckgraefe_colon_NHUC" /tissue_type="colonic mucosa from colon ulcerative colitis patients" /lab_host="DH10B (phage-resistant)" /note="Organ: colon; Vector: pVT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTCGAGCGGCCGCTGCTTTTTTTTTTTT 3'] , double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pVT73 vector. Library went through one round of normalization. Tissue samples were obtained by Dr. Brian Dieckgraefe (Washington University, dieckel@wustl.edu); colonic mucosa represents a range of disease involvement from mild cryptitis to severe ulceration, fibrosis, and degeneration. Library constructed by Bento Soares and M. Fatima Bonaldo."		
	BASE COUNT	166 a	98 c	118 g
	ORIGIN	166 a	98 c	118 g
	Query Match	19.0%	Score 535.2;	DB 10; Length 561;
	Best Local Similarity	99.3%;	Pred. No. 7.2e-123;	
	Matches 537; Conservative	0;	Mismatches 4;	Indels 0; Gaps 0;
	2272	CMCTTCGACCGCGCAATCACGACCTGAAGGGGGAATTACAGGGGCGAGTCTCAT	2331	

REFERENCE 1 (bases 1 to 652)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&st=QV2-CT0261-261099-011-e08&t3=1999-10-26&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 14.
 Location/Qualifiers
 1..652
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0261"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 179 a 153 c 161 g 159 t
 ORIGIN

Query Match 18.5%; Score 520; DB 10; Length 652;
 Best Local Similarity 95.0%; Pred. No. 4.5e-119;
 Matches 570; Conservative 0; Mismatches 25; Indels 5; Gaps 3;

1174 TCCATCTGACGCGCTTCGATGCGGCACTTCTGATAGGAAGAAT--ATCCACATG 1231
 |||||
 53 TCCAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 112
 |||||
 1232 AT--GGATCTGAATTTGCTGCTGACGAGTGGGAGACAACTATTAAGTGGTGGT 1289
 |||||
 113 ATGAGATTCTGTAATTTGCTGCTGACGAGTGGGAGACAACTATTAAGTGGTGGT 172
 |||||
 1290 TAACGAGTCAAAAGTGTGCTCATTCACACAGTGGTGGGCGCTTGC-AG 1348
 |||||
 173 TAACGAGTCAAAAGTGTGCTCATTCACACAGTGGTGGGCGCTTGC-AG 232
 |||||
 1349 CTCAGAGCTGAGAGGTGTCACAAATGACAGAGGTTTACAGCATATGCTTCAAGAT 1408
 |||||
 233 CTCAGAGCTGAGAGGTGTCACAAATGACAGAGGTTTACAGCATATGCTTCAAGAT 292
 |||||
 1409 AAGTTCAGAACATGGCTCATTTGATGCTTTGGGCGCTTTCATCAGAAATGAGAGCG 1468
 |||||
 293 AAGTTCAGAACATGGCTCATTTGATGCTTTGGGCGCTTTCATCAGAAATGAGAGCG 352
 |||||
 1469 TCTCTACAGCTCCATTCAGCTTGAGAGTTAAGGATTAACCTCCAGAACAGCAGTGA 1528
 |||||
 353 TCTCTACAGCTCCATTCAGCTTGAGAGTTAAGGATTAACCTCCAGAACAGCAGTGA 412
 |||||
 1529 TGAATGGACAGTATGCGGAGCAGCAGCAGCAGTGGAAAGGACCTTGTCTTACACT 1588
 |||||
 413 TGAATGGACAGTATGCGGAGCAGCAGCAGCAGTGGAAAGGACCTTGTCTTACACT 472

1589 GGACAGCAGCGCTCCCAATCTTCTGATGATCCAGTGACAGAGCAAGGTGGCT 1648
 |||||
 43 GGACATGACAGCGCTCCCAATCTTCTGATGATCCAGTGACAGAGCAAGGTGGCT 532
 |||||
 169 TTGTAGTGACAAAACACAAATGGCGCTTACCTCCAAATCCAGCATTTGTGAAGTTG 1708
 |||||
 533 TTGTAGTGACAAAACACAAATGGCGCTTACCTCCAAATCCAGCATTTGTGAAGTTG 592
 |||||
 1709 GCACCTGGAATATACAGTCTGACAGCAAGCTTCAACCTTACCTGACTGACAGTCCC 1768
 |||||
 593 GCACCTGGAATATACAGTCTGACAGCAAGCTTCAACCTTACCTGACTGACAGTCCC 652
 |||||

RESULT 13
 LOCUS BF083270 528 bp mRNA EST 18-OCT-2000
 DEFINITION RCI-CT0249-110900-214-b01 CT0249 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF083270
 VERSION BF083270.1 GI:10877100
 KEYWORDS EST.
 SOURCE human;
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 528)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st=RCI-CT0249-110900-214-b01&t3=2000-09-11&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 33
 High quality sequence stop: 526.
 Location/Qualifiers
 1..528
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0249"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 158 a 113 c 127 g 130 t
 ORIGIN

Query Match 18.4%; Score 516.8; DB 11; Length 528;
 Best Local Similarity 98.7%; Pred. No. 2.8e-118;
 Matches 521; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

1940 CCTTGAACTACTGATATGAGCAGGTGTGATGCTACTAAGATGACGCTGTCTACT 1999

Db 1 CGGGATATTTTGGATTAAGAGAGAGGCTTATCTACTAAGAGAGAGGCTCTACT 60
 Oy 2000 CAAGGATTTTACACACTTATGACACGAATGATGATACATGTAAGAGAGGCTCTG 2059
 Db 61 CAAGGATTTTACACACTTATGACACGAATGATGATACATGTAAGAGAGGCTCTG 120
 Oy 2060 GAGGAGTTTAAAGCAGCAGAGAGAGAGTATACCCAGAGAGTGGAGCATGATAC 2119
 Db 121 GAGGAGTTTAAAGCAGCAGAGAGAGTATACCCAGAGAGTGGAGCATGATAC 180
 Oy 2120 CTGGCTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2179
 Db 181 CTGGCTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 Oy 2180 ATGATGTTTAAACAGCAAGAGAGTGTGTTTACAGCAAGATCTGGGAGGCTCTG 2239
 Db 241 ATGATGTTTAAACAGCAAGAGTGTGTTTACAGCAAGATCTGGGAGGCTCTG 300
 Oy 2240 CTGCTGATGTTTCCCAATGCTCCATACCTGATCTCTCCACCTGGCAATACAGC 2299
 Db 301 CTGCTGATGTTTCCCAATGCTCCATACCTGATCTCTCCACCTGGCAATACAGC 360
 Oy 2300 TGAAGCGGGAATTCACGGGGGAGCTCTCATTTATCTGACTTGGAGAGCTCTGG 2359
 Db 361 TGAAGCGGGAATTCACGGGGGAGCTCTCATTTATCTGACTTGGAGAGCTCTGG 420
 Oy 2360 ATTATGACATGAGACAGCTCAGCAAGATGATGATGATGATGATGATGATGATG 2419
 Db 421 ATTATGACATGAGACAGCTCAGCAAGATGATGATGATGATGATGATGATGATG 480
 Oy 2420 TCAGAGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATG 2467
 Db 481 TCAGAGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATG 528

RESULT 14
 AM361523 561 bp mRNA EST 04-FEB-2000
 LOCUS QV2-CT0261-261099-011-905 CT0261 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM361523
 VERSION AM361523.1 GI:6866277
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 561)
 HCGP http://www.ludwig.org.br/ORESTES.
 TITLE The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-CT0261-
 261099-011-905&t3=1999-10-26&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 43
 High quality sequence stop: 560.
 Location/Qualifiers
 1..561

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0261"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site.1: Sma1; Site.2:

Query Match 17.38; Score 487.2; DB 10; Length 561;
 Best Local Similarity 97.9%; Pred. No. 6,9e-111;
 Matches 525; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
 Sma1; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 /716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions.
 BASE COUNT 154 a 140 c 140 g 127 t
 ORIGIN

Oy 1235 GATCTGAATTTGCTCTGTCAC-AGATGGGGAAGACAACTATAGTGGGTCTTAC 1293
 Db 12 GATCTGTAATTTGCTCTGTCACAGGAGATGGGGAAGACAACTATAGTGGGTCTTAC 71
 Oy 1294 GAGCTCAAAACAACTGGTGCATCATCAGACAGCTGCTTTGGGCGCCCTGTC-AGCTCA 1352
 Db 72 GAGCTCATATAGTGGTGCCTCATCAGACAGCTGCTTTGGGCGCCCTGTCAGAGCTCA 131
 Oy 1353 AGAAGTGAAG-AGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTACATCAAG 1411
 Db 132 AGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 191
 Oy 1412 TTGAGACAAATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1471
 Db 142 TTGAGACAAATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 251
 Oy 1472 CTCAGCGCTCCATCCAGCTTGAAGATGATGATGATGATGATGATGATGATGATG 1531
 Db 252 CTCAGCGCTCCATCCAGCTTGAAGATGATGATGATGATGATGATGATGATGATG 311
 Oy 1532 ATGGCACAGTATGCTGAGACAGACCGTGGGAAGACACTTGTCTTATACCTGGA 1591
 Db 312 ATGGCACAGTATGCTGAGACAGACCGTGGGAAGACACTTGTCTTATACCTGGA 371
 Oy 1532 CAAGCGAGCTCCCAATGCTCTCTGAGATGATGATGATGATGATGATGATGATG 1651
 Db 372 CAAGCGAGCTCCCAATGCTCTCTGAGATGATGATGATGATGATGATGATGATG 431
 Oy 1652 TAGTGACAAAACACCAAAATGAGCTTACCTCCAAATCCAGAGCTTGAAGTTGGCA 1711
 Db 432 TAGTGACAAAACACCAAAATGAGCTTACCTCCAAATCCAGAGCTTGAAGTTGGCA 491
 Oy 1712 CTGGAAATACAGCTGTCAGAGCAAGCTTGAACCTGGAACCTGACGTCACGCC 547
 Db 492 CTGGAAATACAGCTGTCAGAGCAAGCTTGAACCTGGAACCTGACGTCACGCC 547

RESULT 15
 A1802693/c 504 bp mRNA EST 18-DEC-1999
 LOCUS W117C01.x1 Soares-Dieckgrafe_colon_NHUC Homo sapiens cDNA clone
 DEFINITION IMAGE:2350848 3' similar to TR:088826 O88826 G08-5 PROTEIN. ; mRNA
 sequence.
 ACCESSION A1802693
 VERSION A1802693.1 GI:5368165
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 504)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaaps-remail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:54:52 ; Search time 521.98 Seconds
(without alignments)
1220.512 Million cell updates/sec

Title: US-09-049-696-18
Perfect score: 2813
Sequence: 1 GAATTCACAGGAGATGTAC.....AAATTAATTCATTCCTTA 2813

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	780.8	27.8	878	1	US-08-469-667-8
2	780.8	27.8	878	5	PCT-US95-07289-8
3	323.8	11.5	401	4	US-09-221-298-34
4	228.2	8.1	576	4	US-09-385-982-23
5	221.4	7.9	595	4	US-09-385-982-25
6	200.8	7.1	618	4	US-09-385-982-24
7	183.4	6.5	611	4	US-09-385-982-27
8	168.6	6.0	742	4	US-09-385-982-33
9	52	1.8	7218	1	US-08-232-463-14
10	38.6	1.4	5156	2	US-09-091-432-3
11	35	1.2	1794	3	US-09-012-515A-13
12	35	1.2	1794	3	US-08-360-144A-13
13	35	1.2	1794	5	PCT-US95-06722-13
14	34.8	1.2	6824	1	US-08-015-973-2
15	34.8	1.2	6824	2	US-08-448-164-2
16	34	1.2	3600	4	US-08-855-910-7
17	34	1.2	5319	4	US-08-169-927-1
18	33.8	1.2	3447	2	US-08-252-995D-3
19	33.8	1.2	3447	2	US-08-834-108-3
20	33.2	1.2	2520	2	US-08-454-557C-50
21	33.2	1.2	2520	2	US-08-340-426D-50
22	33.2	1.2	2520	2	US-08-450-673C-50
23	33.2	1.2	2520	5	PCT-US95-1711A-50
24	33	1.2	1984	1	US-07-885-970A-25
25	33	1.2	1985	1	US-08-298-687A-25
26	33	1.2	1985	1	US-08-298-829-25
27	33	1.2	2615	1	US-08-072-281-1

28	33	1.2	2615	1	US-08-759-446-1	Sequence 1, Appl
29	33	1.2	2615	4	US-09-027-998A-1	Sequence 1, Appl
30	33	1.2	3050	2	US-09-031-442A-21	Sequence 21, Appl
31	33	1.2	3050	4	US-09-258-377-21	Sequence 21, Appl
32	32.8	1.2	1534	1	US-08-300-903A-6	Sequence 6, Appl
33	32.8	1.2	1838	2	US-09-091-432-1	Sequence 1, Appl
34	32.6	1.2	2369	4	US-08-910-925-2	Sequence 2, Appl
35	32.4	1.2	1117	4	US-09-247-373B-33	Sequence 33, Appl
36	32.4	1.2	1423	4	US-08-916-576B-3	Sequence 3, Appl
37	32	1.1	430	4	US-08-905-223-235	Sequence 235, App
38	31.8	1.1	2402	3	US-08-776-265-4	Sequence 4, Appl
39	31.8	1.1	4365	3	US-08-776-265-1	Sequence 1, Appl
40	31.6	1.1	603	4	US-09-411-329C-2	Sequence 2, Appl
41	31.6	1.1	609	4	US-09-411-329C-6	Sequence 6, Appl
42	31.6	1.1	816	4	US-09-411-329C-19	Sequence 19, Appl
43	31.6	1.1	1368	3	US-08-874-563-5	Sequence 5, Appl
44	31.6	1.1	1368	3	US-08-577-483-14	Sequence 14, Appl
45	31.6	1.1	1373	4	US-09-411-329C-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-469-667-8
Sequence 8, Application US/08469667
Patent No. 5733748
GENERAL INFORMATION:

APPLICANT: YU, Guo-Liang

INVENTOR: Rosen, Craig

TITLE OF INVENTION: Colon Specific Genes and Proteins

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carrella, Byrne, Bain, Gillfillan, Cecchi,

ADDRESSEE: Stewart & Olstein

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: NJ

COUNTRY: USA

ZIP: 07068-1739

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,667

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Ferrari, Gregory D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-435

TELEPHONE: 201-994-1744

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 878 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2..685

US-08-469-667-8

Query Match 27.8%; Score 780.8; DB 1; Length 878;
Best Local Similarity 98.9%; Pred. No. 5.1e-231;
Matches 796; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1992 TGTCTACTCAAGGATATTTACAACTTATGACAGCAAGTAGATACAGTGTAAAGTCCG 2051
DB 1 TGTCTACTCAAGGATATTTACAACTTATGACAGCAAGTAGATACAGTGTAAAGTCCG 60
QY 2052 GGCTCTGGAGAGATTAAACGACGACGAGAGAGTATACCCGACGAGTGGAGCT 2111
DB 61 GGCTCTGGAGAGATTAAACGACGACGAGAGAGTATACCCGACGAGTGGAGCT 120
QY 2112 GTACATACCTGGCTGGATTTGAGAAATGATGAATATCAATGAAATCCAGCAAGCTGAAT 2171
DB 121 GTACATACCTGGCTGGATTTGAGAAATGATGAATATCAATGAAATCCAGCAAGCTGAAT 180
QY 2172 TAATAAGATGATGTTCAACAACAAGCAAGTGTGTTTCAGCAGAACATCTCGGAGGCTC 2231
DB 181 TAATAAGATGATGTTCAACAACAAGCAAGTGTGTTTCAGCAGAACATCTCGGAGGCTC 240
QY 2232 ATTTGCTGCTTGTATGTCCTCAAAATGCTCCATACCTGATCTCTTCCCACTGGCCAAAT 2291
DB 241 ATTTGCTGCTTGTATGTCCTCAAAATGCTCCATACCTGATCTCTTCCCACTGGCCAAAT 300
QY 2292 CACGACCTGAAAGGGGAAATTTACGAGGGGAGTCTCATTAATCTGACTTGGACAGCTCC 2351
DB 301 CACGACCTGAAAGGGGAAATTTACGAGGGGAGTCTCATTAATCTGACTTGGACAGCTCC 360
QY 2352 TGGGATGATTTATGACATGGAACAGCTCACAAGTATATCATTCGAATTAAGTACAAGTAT 2411
DB 361 TGGGATGATTTATGACATGGAACAGCTCACAAGTATATCATTCGAATTAAGTACAAGTAT 420
QY 2412 TCTTGATCTCAGACAGTCAATGAAATCTTCAATGAAATCTGCTCATATCC 2471
DB 421 TCTTGATCTCAGACAGTCAATGAAATCTTCAATGAAATCTGCTCATATCC 480
QY 2472 AAAGGAACCAACTCTGAGGAAGCTTTTGTGTTAAACAGAAACATTAATCTTTGAAAA 2531
DB 481 AAAGGAACCAACTCTGAGGAAGCTTTTGTGTTAAACAGAAACATTAATCTTTGAAAA 540
QY 2532 TGGCAGACATCTTTTCATGCTATTCAGCTGCTGATTAAGCTGATGAAATCAGAAAT 2591
DB 541 TGGCAGACATCTTTTCATGCTATTCAGCTGCTGATTAAGCTGATGAAATCAGAAAT 600
QY 2592 ATCCAAATTCAGAGATCTTGTGTTATTCCTCAGACAGCTCCGCAAGACAGCTGAG 2651
DB 601 ATCCAAATTCAGAGATCTTGTGTTATTCCTCAGACAGCTCCGCAAGACAGCTGAG 660
QY 2652 TCGCTGATGAAGCTGCTGCTCTTGT - CCTAATATTCATATCAACAGCAATCTCTGGA 2710
DB 661 TCGCTGATGAAGCTGCTGCTCTTGT - CCTAATATTCATATCAACAGCAATCTCTGGA 720
QY 2711 TTCACATTTTAAAAATTAATGAGGAAGTGAATGAGAACTGCAGCTGTCATATAGCCTAAG 2770
DB 721 TTCACATTTTAAAAATTAATGAGGAAGTGAATGAGAACTGCAGCTGTCATATAGCCTAAG 780
QY 2771 GCTGAATTTTGTGCTGATTAATTAAT 2795
DB 781 GCTGAATTTTGTGCTGATTAATTAAT 805

RESULT 2

PCT-US95-07289-8

Sequence 8, Application PC/TUS9507289

GENERAL INFORMATION:

APPLICANT: Yu, Guo-Liang

APPLICANT: Rosen, Craig

TITLE OF INVENTION: COLON Specific Genes and Proteins

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESS: Carella, Byrne, Bain, Gilfillan, Cecchi,

ADDRESS: Stewart & Olstein

CITY: Roseland

STATE: NJ

COUNTRY: USA

ZIP: 07068-1739

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07289
FILING DATE: 06-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-265
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
FAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
POPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
PCT-US95-07289-8

Query Match 27.8%; Score 780.8; DB 5; Length 878;
Best Local Similarity 98.9%; Pred. No. 5.1e-231;
Matches: 796; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1992 TGTCTACTCAAGGATATTTACAACTTATGACAGCAAGTAGATACAGTGTAAAGTCCG 2051
DB 1 TGTCTACTCAAGGATATTTACAACTTATGACAGCAAGTAGATACAGTGTAAAGTCCG 60
QY 2052 GGCTCTGGAGAGATTAAACGACGACGAGAGAGTATACCCGACGAGTGGAGCT 2111
DB 61 GGCTCTGGAGAGATTAAACGACGACGAGAGAGTATACCCGACGAGTGGAGCT 120
QY 2112 GTACATACCTGGCTGGATTTGAGAAATGATGAATATCAATGAAATCCAGCAAGCTGAAT 2171
DB 121 GTACATACCTGGCTGGATTTGAGAAATGATGAATATCAATGAAATCCAGCAAGCTGAAT 180
QY 2172 TAATAAGATGATGTTCAACAACAAGCAAGTGTGTTTCAGCAGAACATCTCGGAGGCTC 2231
DB 181 TAATAAGATGATGTTCAACAACAAGCAAGTGTGTTTCAGCAGAACATCTCGGAGGCTC 240
QY 2232 ATTTGCTGCTTGTATGTCCTCAAAATGCTCCATACCTGATCTCTTCCCACTGGCCAAAT 2291
DB 241 ATTTGCTGCTTGTATGTCCTCAAAATGCTCCATACCTGATCTCTTCCCACTGGCCAAAT 300
QY 2292 CACGACCTGAAAGGGGAAATTTACGAGGGGAGTCTCATTAATCTGACTTGGACAGCTCC 2351
DB 301 CACGACCTGAAAGGGGAAATTTACGAGGGGAGTCTCATTAATCTGACTTGGACAGCTCC 360
QY 2352 TGGGATGATTTATGACATGGAACAGCTCACAAGTATATCATTCGAATTAAGTACAAGTAT 2411
DB 361 TGGGATGATTTATGACATGGAACAGCTCACAAGTATATCATTCGAATTAAGTACAAGTAT 420
QY 2412 TCTTGATCTCAGACAGTCAATGAAATCTTCAATGAAATCTGCTCATATCC 2471
DB 421 TCTTGATCTCAGACAGTCAATGAAATCTTCAATGAAATCTGCTCATATCC 480
QY 2472 AAAGGAACCAACTCTGAGGAAGCTTTTGTGTTAAACAGAAACATTAATCTTTGAAAA 2531
DB 481 AAAGGAACCAACTCTGAGGAAGCTTTTGTGTTAAACAGAAACATTAATCTTTGAAAA 540
QY 2532 TGGCAGACATCTTTTCATGCTATTCAGCTGCTGATTAAGCTGATGAAATCAGAAAT 2591
DB 541 TGGCAGACATCTTTTCATGCTATTCAGCTGCTGATTAAGCTGATGAAATCAGAAAT 600


```

US-09-385-982-25
: Sequence 25, Application US/09385982
: Patent No. 6262334
:
: GENERAL INFORMATION:
: APPLICANT: ENDEGE, WILSON O., ET AL.
: TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
: TITLE OF INVENTION: PRODUCTS: II
: FILE REFERENCE: CCDNA-260XX
:
: CURRENT APPLICATION NUMBER: US/09/385,982
:
: CURRENT FILING DATE: 1999-08-30
: EARLIER APPLICATION NUMBER: 09/328,111
: EARLIER FILING DATE: 1999-06-08
: EARLIER APPLICATION NUMBER: 60/117,393
: EARLIER FILING DATE: 1999-01-27
: EARLIER APPLICATION NUMBER: 60/098,639
: EARLIER FILING DATE: 1998-08-31
:
: NUMBER OF SEQ ID NOS: 544
:
: SOFTWARE: FastSeq for Windows Version 3.0
:
: SEQ ID NO 25
:
: LENGTH: 595
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(595)
: OTHER INFORMATION: n = A,T,C or G
:
US-09-385-982-25

```

Query Match	Similarity	7.9%	Score 221.4	DB 4	Length 595
Best Local Similarity	68.4%		Pred. No. 2.1e-58		
Matches 355	Conservative	0	Mismatches 155	Indels	9
				Gaps	4
QY	2114	ACATACCTGGCTGGATTGAGAAATGATAATGAATCAATGAAATCCACCCAGACCTGGAATTA	2173		
Db	1	acatacaacgcgtggtagtgaacggggaattgaagcaaacccgcgaagaccatgaattt	60		
QY	2174	ATAAGATGATGTTTAAACACAAAGCAAGTGTGTTTCAGCAAGAAATCTTGGGAGCTCAT	2233		
Db	61	atgaggaatc---tcagaccaccttggaggaattcaacgcgaacagcatccggagtgcat	117		
QY	2234	TTGTGGCTTCATGATCCCAAAATGCTCCCATACCTGATATCTTCCACCTGGCCAAATCA	2293		
Db	118	tttgggtntcaacaaagttcccaagccttccttcctcggaaccaatacccaagaatcaatca	177		
QY	2294	CCGACCTGAAGGCGGAAATTCACGGGGGCGAGTCTTAATATCTGACTTGGACAGCTCTCG	2353		
Db	178	cagaccttgatgcccacagntcatgagg---ataanaattcttcaatgagcaagcccg	234		
QY	2354	GGGATGATTATGACCATGGAACAGCTCACAGATATATCATTTGGAATAGTACAGTATTC	2413		
Db	235	gagataatttgaatgttgyaanaagttcaaacgntatatacataagaatlaagtcgaatctc	294		
QY	2414	TTGATCTCAGAGACAAAGTTCAATGATATCTTCAAGTGAATACTGCTGCATCCCAA	2473		
Db	295	ttgatctaaagaacacgnttntgatgatgctcttcaagtaaatcactgatgctgtaaccaca	354		
QY	2474	AGGAAGCCAACTCTGAGGAAGTCTTTTGTGTTAAACACAGAAAACATATTCTTTTGAANAATG	2533		
Db	355	aggagggccaactcccaangaagcttgcgttttaaacccgaanaatatctcagaagaataatg	414		
QY	2534	GCACAGATCTTTTCATTGCTATTCAGGCTGTTGATTAAGTGCATCTGAAATCAGAAAATAT	2593		
Db	415	caaccacacataattatgctcmtaanaagfatagatanag--caattgacatc--naagtn	471		
QY	2594	CCAACATTCAGAGATATCTTGTGTTATTCGCCACACAGA	2632		
Db	472	tcacacatgnaaagunacttgggtatattccctcagaataa	510		

RESULT 6
US-09-385-982-24/C
; Sequence 24, Application US/09385982

Patent No.6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILED IN: INVENTION: PRODUCTS: II
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 24
LENGTH: 618
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(618)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-24

Query: Match Similarity 7.1%; Score 200.8; DB 4; Length 618;
 Percent Identical 69.5%; Pred. No. 5e-52;
 Matches 303; Conservative 0; Mismatches 127; Indels 6; Gaps 2;

DB	Seq	ACTGCTACTCCGATA	84
DB	59	ACTGCTACTCCGATA	84
DB	2658	GCTCCTTGTCCTAATA	2683
DB	139	ACTTTGTTATTCGCCCAAGCAAAATCTGTGACATTGATCCTACACACTCTCTACTCCT	100
DB	251	GCTATTCAGGCGCTGTGATTAAGSTGCATCTGAAATPAGAAATTCCTCAACATTCGACAGAGA	2610
DB	219	GCAATTAAAGATATGATATAAAGCAATTTGACATCAAAAGTATCCAAACCTTGACACAAGA	160
DB	2611	TCTTTGTTATTCCTCCACAGACTCCGCCAGAGA--CACTAGTCTGATGAAACGCT	2667
DB	231	ATTCAGGGGGGAGCTCATTAATCTGACTTGGAGAGCCCTGGGGATGATATGACAT	2370
DB	456	GTTCAATGAG--ATAGATTATTTCTTACATGACAGCACCGAGGATATTTATGTT	400
DB	2371	GGACAGGCTCACAACTATATCTATTCGAAATTAAGTACAGATATTCCTGATCTCAGACACAAG	2430
DB	309	GGAAAAAGTTCAACGTTATATCATTAAGAAATTAAGTCAAGATATCTGTATCTAGACACAGT	340
DB	2421	TTTCATGATCTCTCTCAAGTGAATCTACTGCTGCATCCCAAGAGGACCAACTGTGAG	2490
DB	389	TTTGATGATGCTCTTCTTCAAGTATATCTCTGATCTGTCAACCAAGAGGCAACTCCAG	280
DB	2421	GAAAGTCTTTTGTGTTAAACACAGAAAAACATTTACTTTGAAAAATGGCAGACTTTTCATT	2550
DB	279	GAAAGCTTTGCAATTTAAACACAGAAAAATATCTCAGAAAGAAATGCAACCCACATATTTATT	220
DB	251	GCTATTCAGGCGCTGTGATTAAGSTGCATCTGAAATPAGAAATTCCTCAACATTCGACAGAGA	2610
DB	219	GCAATTAAAGATATGATATAAAGCAATTTGACATCAAAAGTATCCAAACCTTGACACAAGA	160
DB	2611	TCTTTGTTATTCCTCCACAGACTCCGCCAGAGA--CACTAGTCTGATGAAACGCT	2667
DB	231	ATTCAGGGGGGAGCTCATTAATCTGACTTGGAGAGCCCTGGGGATGATATGACAT	2370
DB	456	GTTCAATGAG--ATAGATTATTTCTTACATGACAGCACCGAGGATATTTATGTT	400
DB	2371	GGACAGGCTCACAACTATATCTATTCGAAATTAAGTACAGATATTCCTGATCTCAGACACAAG	2430
DB	309	GGAAAAAGTTCAACGTTATATCATTAAGAAATTAAGTCAAGATATCTGTATCTAGACACAGT	340
DB	2421	TTTCATGATCTCTCTCAAGTGAATCTACTGCTGCATCCCAAGAGGACCAACTGTGAG	2490
DB	389	TTTGATGATGCTCTTCTTCAAGTATATCTCTGATCTGTCAACCAAGAGGCAACTCCAG	280
DB	2421	GAAAGTCTTTTGTGTTAAACACAGAAAAACATTTACTTTGAAAAATGGCAGACTTTTCATT	2550
DB	279	GAAAGCTTTGCAATTTAAACACAGAAAAATATCTCAGAAAGAAATGCAACCCACATATTTATT	220
DB	251	GCTATTCAGGCGCTGTGATTAAGSTGCATCTGAAATPAGAAATTCCTCAACATTCGACAGAGA	2610
DB	219	GCAATTAAAGATATGATATAAAGCAATTTGACATCAAAAGTATCCAAACCTTGACACAAGA	160
DB	2611	TCTTTGTTATTCCTCCACAGACTCCGCCAGAGA--CACTAGTCTGATGAAACGCT	2667
DB	231	ATTCAGGGGGGAGCTCATTAATCTGACTTGGAGAGCCCTGGGGATGATATGACAT	2370
DB	456	GTTCAATGAG--ATAGATTATTTCTTACATGACAGCACCGAGGATATTTATGTT	400
DB	2371	GGACAGGCTCACAACTATATCTATTCGAAATTAAGTACAGATATTCCTGATCTCAGACACAAG	2430
DB	309	GGAAAAAGTTCAACGTTATATCATTAAGAAATTAAGTCAAGATATCTGTATCTAGACACAGT	340
DB	2421	TTTCATGATCTCTCTCAAGTGAATCTACTGCTGCATCCCAAGAGGACCAACTGTGAG	2490
DB	389	TTTGATGATGCTCTTCTTCAAGTATATCTCTGATCTGTCAACCAAGAGGCAACTCCAG	280
DB	2421	GAAAGTCTTTTGTGTTAAACACAGAAAAACATTTACTTTGAAAAATGGCAGACTTTTCATT	2550
DB	279	GAAAGCTTTGCAATTTAAACACAGAAAAATATCTCAGAAAGAAATGCAACCCACATATTTATT	220
DB	251	GCTATTCAGGCGCTGTGATTAAGSTGCATCTGAAATPAGAAATTCCTCAACATTCGACAGAGA	2610
DB	219	GCAATTAAAGATATGATATAAAGCAATTTGACATCAAAAGTATCCAAACCTTGACACAAGA	160
DB	2611	TCTTTGTTATTCCTCCACAGACTCCGCCAGAGA--CACTAGTCTGATGAAACGCT	2667
DB	231	ATTCAGGGGGGAGCTCATTAATCTGACTTGGAGAGCCCTGGGGATGATATGACAT	2370
DB	456	GTTCAATGAG--ATAGATTATTTCTTACATGACAGCACCGAGGATATTTATGTT	400
DB	2371	GGACAGGCTCACAACTATATCTATTCGAAATTAAGTACAGATATTCCTGATCTCAGACACAAG	2430
DB	309	GGAAAAAGTTCAACGTTATATCATTAAGAAATTAAGTCAAGATATCTGTATCTAGACACAGT	340
DB	2421	TTTCATGATCTCTCTCAAGTGAATCTACTGCTGCATCCCAAGAGGACCAACTGTGAG	2490
DB	389	TTTGATGATGCTCTTCTTCAAGTATATCTCTGATCTGTCAACCAAGAGGCAACTCCAG	280
DB	2421	GAAAGTCTTTTGTGTTAAACACAGAAAAACATTTACTTTGAAAAATGGCAGACTTTTCATT	2550
DB	279	GAAAGCTTTGCAATTTAAACACAGAAAAATATCTCAGAAAGAAATGCAACCCACATATTTATT	220
DB	251	GCTATTCAGGCGCTGTGATTAAGSTGCATCTGAAATPAGAAATTCCTCAACATTCGACAGAGA	2610
DB	219	GCAATTAAAGATATGATATAAAGCAATTTGACATCAAAAGTATCCAAACCTTGACACAAGA	160
DB	2611	TCTTTGTTATTCCTCCACAGACTCCGCCAGAGA--CACTAGTCTGATGAAACGCT	2667
DB	231	ATTCAGGGGGGAGCTCATTAATCTGACTTGGAGAGCCCTGGGGATGATATGACAT	2370
DB	456	GTTCAATGAG--ATAGATTATTTCTTACATGACAGCACCGAGGATATTTATGTT	400
DB	2371	GGACAGGCTCACAACTATATCTATTCGAAATTAAGTACAGATATTCCTGATCTCAGACACAAG	2430
DB	309	GGAAAAAGTTCAACGTTATATCATTAAGAAATTAAGTCAAGATATCTGTATCTAGACACAGT	340
DB	2421	TTTCATGATCTCTCTCAAGTGAATCTACTGCTGCATCCCAAGAGGACCAACTGTGAG	2490
DB	389	TTTGATGATGCTCTTCTTCAAGTATATCTCTGATCTGTCAACCAAGAGGCAACTCCAG	280
DB	2421	GAAAGTCTTTTGTGTTAAACACAGAAAAACATTTACTTTGAAAAATGGCAGACTTTTCATT	2550
DB	279	GAAAGCTTTGCAATTTAAACACAGAAAAATATCTCAGAAAGAAATGCAACCCACATAT	

RESULT: 17
 US-09-335-982-27
 Sequence: 27 Application US/09385982
 Patent No. 6262334
 GENERAL INFORMATION:
 APPLICANT: ENDIGE, WILSON O., ET AL.
 TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 TITLE OF INVENTION: PRODUCTS: II
 FILE REFERENCE: CCDNA-2603X

RESULT 8
US-09-385-982-33
Sequence 33, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS: II
FILE REFERENCE: CCDA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

APPLICANT: Schlessinger, Joseph
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
TITLE OF INVENTION: PHOSPHATASE-BETA
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/015,973
FILING DATE: 10-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6924 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6924
US-08-015-973-2

Query Match 1.2%; Score 34.8; DB 1; Length 6924;
Best Local Similarity 48.5%; Pred. No. 4.1;
Matches 96; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 53 TCTGATTCCTCACCCTCTAGAGGGGCCGAGTAATTCACCTCACTGAGCTGACACACA 112
DB 5927 TCTTATTCATGATACACTGTTGAGGCCATCTAGTAAGAAACTGAGGCTGCGACA 5986
QY 113 ATGGCTATGAAGGCAATGTCGTCATCGACCCCAATGTGCCAGAGATGAACACTCA 172
DB 5987 GTCATATTCATGCCATGTTATATGCACTCCATTCCTGAGACACAGCAAAACAAAGC 6046
QY 173 TTCAACAATAAAGACATGGTGACCCGAGCATCTCTGTATCTGTTGAAGCTACAGAA 232
DB 6047 TAGAAGAAACAATTCGAGCTCCGACCGAGTCAAAATATACAGAGAGTCAATTCGAG 6106
QY 233 AGCGATTTTATTTCAAA 250
DB 6107 CCTTAAGCAATGCAACA 6124

RESULT 15
US-08-448-164-2
Sequence 2, Application US/08448164
Patent No. 5925536

GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
TITLE OF INVENTION: PHOSPHATASE-BETA
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,164
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/015,973
FILING DATE: 10-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6924 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6924
US-08-448-164-2

Query Match 1.2%; Score 34.8; DB 2; Length 6924;
Best Local Similarity 48.5%; Pred. No. 4.1;
Matches 96; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 53 TCTGATTCCTCACCCTCTAGAGGGGCCGAGTAATTCACCTCACTGAGCTGACACACA 112
DB 5927 TCTTATTCATGATACACTGTTGAGGCCATCTAGTAAGAAACTGAGGCTGCGACA 5986
QY 113 ATGGCTATGAAGGCAATGTCGTCATCGACCCCAATGTGCCAGAGATGAACACTCA 172
DB 5987 GTCATATTCATGCCATGTTATATGCACTCCATTCCTGAGACAGCAAAACAAAGC 6046
QY 173 TTCAACAATAAAGACATGGTGACCCGAGCATCTCTGTATCTGTTGAAGCTACAGAA 232
DB 6047 TAGAAGAAACAATTCGAGCTCCGACCGAGTCAAAATATACAGAGAGTCAATTCGAG 6106
QY 233 AGCGATTTTATTTCAAA 250
DB 6107 CCTTAAGCAATGCAACA 6124

Search completed: April 3, 2002, 20:55:58
Job time: 41916 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:18:34 ; Search time 1321.64 Seconds
(without alignments)
1824.745 Million cell updates/sec

Title: US-09-049-696-18
Perfect score: 2813
Sequence: 1 GAATCAGGAGATGATC.....AAATTAATCATTCCTTA 2813

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq.1101.*
1: /SIDSI/gcgdata/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2812	100.0	2854	22	AAH34879 Human colon cancer
2	2812	100.0	2854	22	AAH81787 Human secreted pro
3	2812	100.0	3111	20	AAZ09840 Human membrane spa
4	2808.4	99.8	2825	22	AAH46124 Human C1CA1 CDNA,
5	2807.2	99.8	3311	22	AAI29502 C902P determined c
6	2797.8	99.5	2867	22	AAH33285 Human colon cancer
7	2743.4	97.4	2745	20	AAH81927 Human ICACC-1 nucl
8	2738.8	97.4	2742	22	AAH46102 Human C1CA1 coding
9	2622.6	93.2	3109	22	AAH35019 Human colon cancer
10	1743	62.0	2843	22	AAH46120 Mouse Gob-5 CDNA,
11	1738.2	61.8	2931	20	AAH81925 Murine ICACC-1 nuc

12	1733	61.6	2739	22	AAH46101	Mouse Gob-5 coding
13	1304	46.4	3265	21	AAZ65095	Membrane-bound pro
14	1304	46.4	3265	22	AAH92092	Human PRO1124 CDNA
15	1304	46.4	3265	22	AAH44241	Human PRO1124 (UNO
16	1022.8	36.4	2616	21	AAH46435	Clone 2516888 of a
17	780.8	27.8	878	18	AAH45884	Human colon specif
18	780.8	27.8	878	19	AAH16672	Polynucleotide seq
19	704.4	25.0	1802	21	AAH98067	Human colon cancer
20	704.4	25.0	1802	22	AAH33192	Human colon cancer
21	554.6	19.7	2784	21	AAH24658	Human lung tumor a
22	554.6	19.7	2784	21	AAH24658	Human lung cancer a
23	552.2	19.6	3951	20	AAH24653	Human lung tumor a
24	552.2	19.6	3951	21	AAH24657	Human lung cancer-
25	552.2	18.9	8031	21	AAH251625	Human membrane cha
26	531.4	18.9	3156	21	AAH81926	Human ICACC-2 nucl
27	512.6	18.2	3190	20	AAH81926	Human tumor-associ
28	433.8	15.4	2454	22	AAH24657	Human lung cancer-
29	366	13.0	3362	21	AAH24657	Human lung tumor a
30	364.4	13.0	3362	22	AAH24657	Human Gob-5 CDNA h
31	326	11.6	486	22	AAH46121	CDNA encoding huma
32	323.8	11.5	401	21	AAH7755	Colon tumour relat
33	323.8	11.5	401	22	AAH7755	Human gene signatu
34	223.2	8.2	255	16	AAH22483	Human colon cancer
35	228.2	8.1	576	21	AAH16018	Human colon cancer
36	221.4	7.9	595	21	AAH16020	Human colon cancer
37	200.8	7.1	618	21	AAH16019	Human colon cancer
38	183.4	6.5	611	21	AAH16022	Human colon cancer
39	168.6	6.0	742	21	AAH16028	Novel human polyn
40	151.6	5.4	375	22	AAH5887	Human lung tumor a
41	128.8	4.6	585	20	AAH24629	Human lung cancer-
42	128.8	4.6	590	20	AAH24626	Human lung tumor a
43	124.2	4.4	546	21	AAH5885	Human lung cancer-
44	124.2	4.4	546	21	AAH5885	Human lung cancer-
45	91.2	3.2	350	22	AAH72429	Human cervical can

ALIGNMENTS

RESULT 1
ID AAH34879 standard; CDNA: 2854 BP.
AAH34879:
AC
XX
DT 03-SEP-2001 (first entry)
XX
XX
XX Human colon cancer antigen encoding CDNA SEQ ID NO:1961.
XX
XX Human colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; chromosome 1; ss.
XX
XX Homo sapiens:
XX
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000MO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
XX
XX 03-NOV-1999; 99US-0163280.
XX
XX (HWA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI: 2001-235357/24.
XX
XX F-PSDB; AAG5474.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX
XX useful for preventing, diagnosing and/or treating colorectal cancers -

PS Claim 1: Page 3462-3463; 9803pp; English.
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX

SO Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match 100.0%; Score 2812; DB 22; Length 2854;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2812; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAATCACAGGAGATGTACAGCAATGGGCCATTAAAGATGTCGTTCATCTTGATT 60
DB 11 gaaatcacaggagatgtacagcaatgggccattaaagatgtctgttcattcttgatt 70
OY 61 CTTCCTCTAGAGAGGGCCCTGATTAATCACTACCTACCTGAGCTGAGCAACAAATGGCAT 120
DB 71 ctctcctctagagaggccctgatttaattcaactacctgagctgagcaacaatggcat 130
OY 121 GAAGGATGTGCTGTAATGCAGCCCAATGTGCCAGAAATGAACAACATCAATCAACA 180
DB 131 gaaggatgtgctgtaatgcagcccaatgtgccagaaatgaacaacatcaatcaaca 190
OY 181 ATAAGGACATGCTGATACCCAGCATCTGTATCTGTTGAAGCTACAGGAAAGGATTT 240
DB 191 ataaggacatgctgataccagcatctgtatctgttgaagctacaggaaaggattt 250
OY 241 TATTTCAAAAATGTGCAATTTGATCTGTAACATGAGAAAGCAAGGCTGATGTCG 300
DB 251 tatttcaaaaatgtgcattttgattctgtaaacatgagaaagcaaggctgattgctg 310
OY 301 AGACCAAACTTGAGACCTACAAAATGCTGATGTTGCTGCTGAGCTTACTCTCCA 360
DB 311 agaccaaaacttgagacctacaaaatgctgattgttgcctgagcttactctcca 370
OY 361 GGTATGATGAACCTACACTGAGAGATGGCACTGTGGAGAAAGGTTAAAGATC 420
DB 371 ggtatgatgaacctacactgagagatggcactgtggagaaaggttaaagatc 430
OY 421 CACCTCACTCTGATTTGATTCAGAGAAAAGTTAGCTGAATGAGCAACAAGTGG 480
DB 431 cactctactctgatttcatctgagaaaagtttagctgaatgagcaacaagtgg 490
OY 481 GCATTTGTCATGAGTGGCTCATCTACCATGGGAGATTTGAGACAGTACAAATGAT 540
DB 491 gcatTTGTCATGAGTGGCTCATCTACCATGGGAGATTTGAGACAGTACAAATGAT 550
OY 541 GAGAAATCTTACTTATCCAAATGGAAGATACAGCAATGATGTTCCAGCAGTTTACT 600
DB 551 gagaatcttacttatactcaaatggagaaatacaagcagtaagtgttcagcaggttact 610
OY 601 GGTACAAATGTAGTAAAGAGTGTACAGGAGGACGCTGTATACCAAAAAGATGACATTC 660
DB 611 ggtacaatgtagtaaaagagtgtacaggagagcgtgtataccaaaaagatgacattc 670

OY 661 AATAAGTAAACGACACTATGAAAAAGAGATGTGTTCTTCCAAATCCCGCAGACG 720
DB 671 aataagtaaacgacctatgaaaaagagatgtgttcttccaaatcccgccagacg 730
OY 721 GAGAAAGCTTTATTAATGTTTTCACACAAATGTTGATTAATGTAATTCGTACAGAA 780
DB 731 gagaagctttcataatgttgcacaacaatgttgcataatgttgcataatcctgcagaa 790
OY 791 CAAAACCCAAACAAAGAGCTCCAAACAAAGCAAAAATGCAAAATCCGAAAGCA 840
DB 851 caaaacccaacaaagagctccaaacaaagcaaaaatgcaaaaatcccgaaagca 850
OY 841 TGGCAAGTAAATCCGATTTGAGAGCTTTAAAGAAACCACTCTATATACAAACAGCA 900
DB 851 tggcaagttaaatccgatttgagagctttaaagaaacctctatatacacaagaagca 910
OY 901 CCAATTCACACCTTCTCATTTGCTGAGATTGAGCAAAAGATTTGTTAAGTCTTGAC 960
DB 911 ccaatccacctctcatcttgcagattgacaagaattgtgtttagtcttgac 970
OY 961 AAATGTGAGATGAGGAGCTGTAAACCGCTCAATGCAATGCAATGCAAGCAGCTT 1020
DB 971 aaatcgtgagatgagagctgtaaacgctcaatgcaatgcaatgcaagcagctt 1030
OY 1021 TTCTGCTGACAGACAGTGTGAGCTGGGCTCTGGATTGGAGTGTGACATTTGACAGTCT 1080
DB 1031 ttctgctgacagacagtgtagctgggctctggattggagtgtagcttgaacgtgct 1090
OY 1081 GCCCATGTCAAAAGTGAACCTCATACAGATTAACAGTGGCAGTGCAGACACACTGCCC 1140
DB 1091 gcccatgtcaaaagtgaacctcatagataaacagtgtagtgaaggaacacactgcc 1150
OY 1141 AAAAGATACCGTCGACGAGCTTCAGAGAGAGAGCTCACTGAGGAGGCTTCGATCGGCA 1200
DB 1151 aaaagatacctgcagcagcttcaagaggagagctcagagcggcttcgacgagca 1210
OY 1201 TTTACTGTATTAGAGAAATATCACTGATGATGATGTAATTTGCTGCTGACGAGAT 1260
DB 1211 ttactgtattagagaaatattcaactgattgataatgtgtcgtgacgagat 1270
OY 1261 GGGGAAGAACACATATTAAGTGGTCTTTAAAGAGGTCAAAAGTGGTCCATATC 1320
DB 1271 ggggaagaacacatattaaagtggcttttaagaggtcaaaagtggctccatattc 1330
OY 1321 CACACAGTGCCTTTGGGCGCTCTGACACTCAAGAACTAGAGAGACTCTCCAAAATGACA 1380
DB 1331 cacacagtgcctttgggcgctctgacctcaagaaactagagagactgctccaaaatgaca 1390
OY 1381 GGAGTTTACAGACATATGCTTCAGATCAAGTTGCAACAATGGCTGATTTGATGCTTTT 1440
DB 1391 ggaagtttacagacatattgcttcagatcaagtttcagacaatggctgattgattt 1450
OY 1441 GGGGCCCTTTATCATCGAAGAAATGAGACTGTCTCAGCGCTCATCTACACTTGAAGTAA 1500
DB 1451 gggggccctttatcatcgaagaaatgagactgtctcagcgctcatcttgaagttaa 1510
OY 1501 GGATTAACCTTCAGAAACAGCAGTGTGATGATGCAAGTATGTTGAGACAGCAGCTG 1560
DB 1511 ggaatTAACCTTCAGAAACAGCAGTGTGATGATGCAAGTATGTTGAGACAGCAGCTG 1570
OY 1561 GGAAGGACACTTTGTTCTTATCACTGACCAACGACGCTCCCAAAATCCTTCTG 1620
DB 1571 ggaaggacactttgttcttataccctggaacagcagctcccaaaatccttctcgg 1630
OY 1621 GATCCCATGAGACAAACAGTGTGCTTTGATGAGCAAAAACCAAAATGGCTTAC 1680
DB 1631 gatcccatgagacaaacagtggttgaagcaaaaaccaaataatggcttacc 1690
OY 1681 CTCCAATCCAGAGGATTTGCAAGTGTGCACTTGAATATCAGTCTCAAGCAAGCTCA 1740
DB 1691 ctccaatccagagattgtcaagtgtgcaacttggaatcagctcagcaagagctcca 1750
OY 1741 CAAACCTTGACCTGACGTGTCACGTCCTGCGTGTCACAAATGCTACCTGCTTCAATTACA 1800

Db	1751	caaaccttgacccttgctgctgcaagctcccgctggtccaaatgctacccttgctccaaattaca	1810
Qy	1801	GTGACTTCCAAAACGACAGACACACGCAAAATGCCAGCCCTCTGTGATTTATGCA	1860
Db	1811	gtgacttccaaaacgacaagagacacaaacaaattcccagcccttgtagttatgca	1870
Qy	1861	AATATTGGCCAAAGAGACTGCCAAATTTCCAGGGCAGGTGCAGAGCCCTGATTGAATCA	1920
Db	1871	aattctgcgaagagagctcccccattccagggccagagtgctcaagccctgattgaaaca	1930
Qy	1921	GTGAATGGAAAAACAGTTACTCTTGGAATCTAGTAAATGAGACAGGTCTGATGCTACT	1980
Db	1931	gtgagtggaaaaaacagttaccttggaactctggaataatgagcgaggtgctgagtact	1990
Qy	1981	AAGGATACGCTGTCTATCCCAAGGTATTTTCACAACTTATGACACCAATGATGATACGT	2040
Db	1991	aaggaatgacggtgtctactcaaggtatttcaaacattgacacgaaatgtagatacagt	2050
Qy	2041	GTAAGATGGGGGCTCTGGGAGAGATTAACGACAGCAGACGAGAGATGATACCCACAG	2100
Db	2051	gtaaagtgagggtctctgggaagagttaacgacgacgacgagcgaggtgatacccaag	2110
Qy	2101	AGTGGAGACTGTATCACTACCTGGCTGTGATTGAAATGATGAATTAATGAAATCCACCA	2160
Db	2111	agtggagcaactgtacatacctggtgatttggaatgataatgaaatcaatggaattcacca	2170
Qy	2161	AGACCTGAATTAATPAAGATGATGTTCAACCAAGCAAGTGTGTTAGCAGAAATACC	2220
Db	2171	agacactgaaattaaatagatgagtgttccaacaaagcaagtggtgttcaagaagaatccc	2230
Qy	2221	TCGGGAGGCTCATTTTGTGCTTCTGTGATGTCGCCAAATGTCCCATACCGATCTTCCCA	2280
Db	2231	tcgggaggtcattcttggtctctgtgtgtcccaatgtcccaataccgtactcttccca	2290
Qy	2281	CTGTGGCCAAATCAACGACTGGAAGGGGAAATTCACGGGGGACGTCTCATTAATCTGACT	2340
Db	2291	ctgtggccaaatccacgaccctggaagggggaattccagggggaggtctaatattctgact	2350
Qy	2341	TGGACAGCTTCTGGGATGATTTATGACCAATGGAAACGCTACACAAATATATCTGAAAT	2400
Db	2351	tggacagctctcctgggagtattatgaccatlgbaaacgctccaaagtatatcatctcgaa	2410
Qy	2401	AGTACAAATTTTGTATCTCAGAGCAAGTTCAATGAATCTCTTCAGATGATACTACT	2460
Db	2411	agtaacaagtattcttcatctcagagagaagttcaatgaatctcttccaagtgaatacact	2470
Qy	2461	GCCTCATCCCAAGAGAGCAACTGTGAGGAAGCTTTTGTATTAACCGAAGAAACATT	2520
Db	2471	gcctcatcccaagaagaagccaactctgggaagctcttctgtttaaacccgaaaaact	2530
Qy	2521	ACTTTTGAAGAAATGGCAGAGATCTTTTCATTTGATTCAGGCTGTGATPAAGTGCATCTG	2580
Db	2531	acttttgaagaaatgacagactcttctcatcttgctatctcagctgctgtgataagtgactg	2590
Qy	2581	AAATTCGAATTAATCCAACTGTGACAGATATCTTTGTATTTCTCCACAGACTCCGCA	2640
Db	2591	aaatccgaatataccaacatlgacgagatattctgttatctccccaagaactccgcga	2650
Qy	2641	GAGACACTATGATCTCATGTAAGAAAGTGGCTCTGTGCTCAATATATCATATCAACGAC	2700
Db	2651	gagacactagctcctgagtgaagaagctgtctctgtctccaaatcatcatcaacagcacc	2710
Qy	2701	ATTTCCTGGCACTTACATTTTAAAAATTATGTGGAAGTGGATGAGAACTGCAGCTGCA	2760
Db	2711	atttcctggcattcacattttaaaaattatgtagaagtgatagaggaaactgcagctgtca	2770
Qy	2761	ATAGCCCTAGGCTGGAATTTTGTGCATATAATTAATTAATCATCATCTCTT	2812
Db	2771	atagccctagggtcgaattcttgtcagataataaataatcatcatccctt	2822
RESULT	2		

AA81787
ID AA81787 standard; cDNA: 2854 BP.
AC AA81787;
XX 12-JUN-2001 (first entry)
DE Human secreted protein gene 1 SEQ ID NO:11.
XX
KW Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
KW dermatological; immunosuppressive; anti-infectious; anti-HIV;
KW immunostimulant; cytosolic; cardiac; vascular; anti-angiogenic;
KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
KW arizalzheimer; antiparkinsonian; antimicrobial; vulnary; gene therapy;
KW immune disorder; hyperproliferative disorder; cardiovascular disease;
KW cancer; angiogenic disorder; neurological disorder; infectious disease;
KW wound healing; regeneration; chemotaxis; chromosome 1; ss.
XX Homo sapiens;
XX NC_00112775-A2.
FD PD
XX 22-FEB-2001.
XX
XX 16-AUG-2000; 2000WO-US223235.
PE 17-AUG-1999; 99US-0149182.
PR 17-AUG-1999; 99US-0149182.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;
PI Blase CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;
PI
XX WP: 2001-147550/15.
DR P-PSDB; AAB74733.
XX
XX Nucleic acids encoding 25 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
PT
XX
XX Claim 1; Page 441; 485pp; English.

AA81787 to AA81817 encode the human secreted proteins given in AAB74733
to AAB74772. Human secreted proteins can have activities based on the
tissues and cells they are expressed in. Example of activities include:
immunomodulatory; antisclerotic; dermatological; immunosuppressive;
anti-inflammatory; anti-HIV; immunostimulant; cytotoxic; cardiac;
vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial; and
vulnerable. Human secreted proteins can be used in gene therapy and
vaccine. Human secreted protein nucleotide sequences (NMI) and proteins
(PPI) may be used in the prevention, diagnosis and treatment of diseases
associated with inappropriate polypeptide expression. For example, NMI
and PPI may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patients genome
that affect the activity of proteins by expressing inactive proteins or
CC to supplement the patients own production of polypeptides. Disorders that
CC may be prevented, diagnosed and/or treated include immune disorders,
CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
CC autoimmune disorders, neurological disorders, infectious diseases and/or
CC fo. promoting wound healing, regeneration and /or chemotaxis. AAB74778 to
CC AAB74786 and AAB74732 represent sequences used in the exemplification of
CC the present invention.

Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match 100.0%; Score 2812; DB 22; Length 2854;
Best Local Similarity 100.0%; Prid. No. 0;
Matches 2812; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCACGAGCAGTGTACAGCATGGGCCATTTAAGAGTTCTGTTCATCTTGATT 60

[illegible]

QY	1141	AAAGATTACTGACGACGCTTACGAGGAGACGCTCCATCTGCACGGGCGCTTGATGCGCA	1200
Db	1137	aaaagattacctgacgacgaccttcaagggagcgtccatctgcacgcyggcttcgacitcgca	1210
QY	1209	TTTACTGTGATTAGGAAGAAATATCCAACTGATGAGATCTGAAATTGCTGTGCAGGAT	1260
Db	1211	tttaactggtatlaggaagaaatatccaactgtagatctcgaatatgtgctgctgcagat	1270
QY	1261	GGGGAGACAACACTATTAAGTGGGTGCTTTAACGAGTCAAAACAAGTGTGTCATATC	1320
Db	1271	ggggagagacaacactataaagtggctttaaagagtcataacaagaatggtgcatactc	1330
QY	1321	CACACAGTCGGTTTGGGGCCCTCGCAGCTCAAGAACTGAGAGCTGTGCCAAATGACA	1380
Db	1331	cacacagtcggtttggggccctcgcagcctcaagaactagggagctgtccaaatgaca	1390
QY	1381	GGAGGTTTACGACATATGCTTCAGATCAATGATCAACAAATAGGCTCATTTGATCTTTT	1440
Db	1391	ggaggtttacgacataatgcttcaagatcaagttcgaacaatggtcctatgtatctt	1450
QY	1441	GGGGCCCTTTTATAGCAAAATGGAAGCTGTCTTCACGCTCCATCCAGCTTGAGATAG	1500
Db	1451	ggggccctttatacgaagaaatggagctgtcttcacgctccatccagcttgaagtag	1510
QY	1501	GGATTAAACCCCTCCAGAACACACAGTGGATTAATGAGCAGATGATGGAGCAGCACCGTG	1560
Db	1511	ggattaaacccctccagaaacagccagctgtagatggagcaagtgatcgttgacagcacgtg	1570
QY	1561	GGAAAGACACTTTGTTTCTTATACCTGACACAGCAGCTCCCAAAATCCTTCTGTGG	1620
Db	1571	ggaaagacactttgtttcttataccttggacaagcagctcccaaatccctctctgg	1630
QY	1621	GATCCCAATGACACGACGACAAAGGCGCTTGTGTGTGACAAAACACAAAATGGCTTAC	1680
Db	1631	gatcccaatgacacgacgacaaaggcgcttgtgtgtgacaaaaacacaaaatggcctac	1690
QY	1681	CTCCAAATCCCAAGGCAATGTGCTAAGGTGGCACTTGGAAATACAGTCTGCACAGCACTCA	1740
Db	1691	ctccaaatcccaaggcatctgttaaggttgcacttggaaatacagctcgtgaagcaagctca	1750
QY	1741	CAAACTTGACCCCGACTGTGCAGTCCCGCGCTCCAAATGTAACCTCGTCCCAATTATCA	1800
Db	1751	caaaacttgaccccgactgtgcacgtcccgctggctcaaatgctacccgtcccaattaca	1810
QY	1811	GTGACTTCCAAAAGCAACAAGGACACACAGCAATTTCCCGAGCCCTGTGTGATTATGCA	1860
Db	1811	gtgacttccaaaagcaacaagagcacacagaaattcccaagccctctgtgttatgca	1870
QY	1861	AATATTGGCCAAAGAGCTCTCCCAATTTTCAGGGCCAGTGTACAGCCCTGATTGAATCA	1920
Db	1871	aatatctgcgaagagagctccccaattcttaagggccaagtgtaacagccctgattgaatca	1930
QY	1921	GTGAATGGAAAAACAGTTACTCTTGAAACTGTGAAATAGTGAAGAGCAGGTGTGATGACTCT	1980
Db	1931	gtgaatggaaaaaacagtlactcttggaaactacgtgataataggagccaggtgctgcagctact	1990
QY	1981	AAGATGACGGTGTCTACTCAAGGTATTTTCAACAATTATGACACGAATGTAGATACAGT	2040
Db	1991	aagatgacggtgtctactcaagtlatttccaactttagacagaaatggtatagtaagtt	2050
QY	2041	GTTAAATCGGGGCGCTGTGGAGAGATTAAACGACGACGAGAGGTATATCCACAGAG	2100
Db	2051	gttaaatcgggcgctgtggagagattaaacgacgacgagagcagagagtgatacccaagag	2110
QY	2101	AGTGGAGACTGTATACATCTGCTGCTGATTTGAAATGATGAATTCATATGAAATCCACA	2160
Db	2111	agtggagactgtatatactctgctgatttggaaatgataatacaatgtgaatccaca	2170
QY	2161	AGACCTCAAAATTAATGAAGATGATGTTTCAACACACAAAGTGTGTTTACGACGAATCC	2220
Db	2171	agaccctcaaaatataaagatgtatgttcaacaaagcaagtggttcaagaagaatcc	2230

QY	2221	TCGGGAGGCTGATTGTGGCTTCTTGATGTGCCAAATCTCCCAATCTGATCTTCCCA	2280
Db	2221	tcgggagggctcatttgygctcttcgagtcccaatgctcccaactgactgctcttccca	2290
QY	2281	CCTGGCCAAATCACCGACCTGTAGAGCGGAAATTCCAGCGGGGAGTCTCATTTAATGACT	2340
Db	2291	cttgcgccaatcacgcgacctgaaggcgaaattcaacggggcgactccatcaattacgtact	2350
QY	2341	TGSGACGCTCTGGGGMATGATTATGACCATGGAACGACTCCACAGTATATCATTCGATA	2400
Db	2351	tgsgaacgctctcggggagatlaagaccatggaacgacgctccaaaglatatcaltcgata	2410
QY	2401	AGTACAAGTATTTCTTGATCTCAGAGACAAAGTTCATGATCTCTTCAAGTGAATACTACT	2460
Db	2411	agtaacaagtattcttgatctcagagacaaggtcaatgaatctctctcaagtgaaatactact	2470
QY	2461	GCTCTCATCCCAAGGAAAGCCACTCTGTAGAGAACTCTTTTGTTTAAACGAAACAATT	2520
Db	2471	gctctcatccccaagaagcaactctgaggaagctcttltgtltaaacccgaaaaacatt	2530
QY	2521	ACTTTTGAAATGCGACAGATCTTTTCAATTGCTATTCAGGCTGTGTGAATAGTGCATCTG	2580
Db	2531	acttttgaatatgtgcacgacatcttcttcattgtctatccaagctgttgataagtgatcgtg	2590
QY	2581	AAATCAGAAATATCCACATTTGCACGAGTATCTTTGTTTATTCCTCCACAGACTCCGSCCA	2640
Db	2591	aaatcagaataatccaaacattgcacgaglatctltgtltaactccacacagactcgcga	2650
QY	2641	GAGAACCTAGTCCGTGATGAAAGCTCTGCTCTCTCTATATTTCATATCAACAGACACC	2700
Db	2651	gagaaacctagctcgtatgaaacgctcgtcctctgtccataatcatatcaaacagacc	2710
QY	2701	ATTCTGTGCATTTCACATTTTAAAAATTATGTGGAAGTGGATAGGAGAACTGCACGTGTCA	2760
Db	2711	attctgtgcattcacattttaaaaaattatgtggaagtgataggaagaaactgcagctgtca	2770
QY	2761	ATAGGCTTAGGGCTGAATTTTGTGCAGATTAATAATAATTAATCATTCATCCTT	2812
Db	2771	atagccttagggctgaattttgtcagataataataaataatcatctacacctt	2822
RESULT 3			
AAZ09840			
ID	AAZ09840	standard; cDNA; 3111 BP.	
XX	AAZ09840;		
AC			
XX			
DT	26-NOV-1999	(first entry)	
XX			
DE	Human membrane spanning protein MSP-5 cDNA fragment 2.		
XX			
KW	Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;		
KW	neoplastic disorder; immunological disorder; reproductive disorder;		
KW	MSP-5; ds.		
XX			
OS	Homo sapiens.		
XX			
PM	W09946380-A2.		
XX			
PD	16-SEP-1999.		
XX			
PF	09-MAR-1999; 99WO-US05073.		
XX			
PR	13-MAR-1998; 98US-0039064.		
XX			
PA	(INCY-) INCYTE PHARM INC.		
XX			
PI	Tang YT, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;		
XX	Guegler KJ, Kaser MR, Baughn MR, Shah P.		
XX	WPI: 1999-551409/46.		
DR	P-PSDB; AAY33298.		
XX			

PT	Neg. human membrane spanning proteins used to, e.g. prevent and treat neoplastic disorders -
PS	Example 1: Page 80-81; 81bp; English.
AA	This invention describes novel human membrane spanning proteins (MSPs),
CC	and the polynucleotides encoding them. The products of the invention are
GC	used to diagnose, prevent and treat neoplastic, immunological and
CC	reproductive disorders. This sequence encodes a human membrane spanning
CC	protein MSP-3 fragment.
AA	Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;
EE	
Query Match	100.0%; Score 2812; DB 20; Length 3111;
Best Local Similarity	100.0%; Pred. NO. 0;
Matches 2812:	Conservative 0; Mismatches 0; Indels 0; Gaps 0
CY	GAAATCAGAGGAGATGTACACCAATGGGGCCATTAAAGTTCTCTTCATCTTGATT 60
DP	gaatcacagaggagtgtacagcaatgggcccatttaagaagtcttgttcatccttgatt 69
CF	CTTACCCTTTGAAGGGGCCCTGAGTAATTACTCATTCACGTACAACAATGGCTAT 120
DF	cctccacctctgaagaggggccctgagtaattaccatttcagctgacaacaatggtat 129
OY	GAAGCATGTGCTGTGCATCGACCCCAATGTGCCAGAAGATGAANAACCTCATTCANCA 180
DO	gaagcaltgtcgtttgcacatgcacccaatgtgccagaagaatgaaacactcattcaaaa 189
OY	ATAAAGACATGNTGACCCAGGCATCTCTGTATCTGTTTGAAGCTACAGAAAGCGATT 240
DO	ataaagacalbtgtaccacagcatctctgtatcttgtttaagcttacaggaagcgatt 249
OY	TATTTCAAATGTTGCCATTGTGATTTCTGAAATGGAABACAAAGGCTACATGTG 300
DO	tattcaaaaaagtgtgcatttgcatttccgaaacaatggaagaaagctgcactatgtg 309
OY	AGACCAAAATTTGAGACCTTACAAAATGCTGATGTTCTGTTGCTGAGTCTACTCTCCA 360
DO	agaccaaaactltgagacctaacaaaaatgctgtgttctgttgcgtgactactctcca 369
OY	GSTAAATGATGAACCTTACACTGACAGATGGCGAACCTGTGGAGAGAGAGGCTGAAGATC 420
DO	gstaaatgtgaaccttaccactcagtagatgtggcgaacttggagagaggtgaaagattc 429
OY	CACCATCTCCGATTTATTCAGAGAAAAAGTTAGCTGAATFATGACACACAAGGTAG 480
DO	caacctcattcctgatattcatgtcagggaaaaaatgttagctgaatatgagccacaagttag 489
OY	GCATTGTCCATGAGTGGGCTCATCTACGATGGGAGTAGTATTGACGAGTACAAATAATGAT 540
DO	gcatttgcacatgagcttgggctcatctacagtcagtgaggatatttgcagcagtaacaataatgatt 549
OY	GAGAAATCTACTATTCCAAATGAGAAATACAAAGAGTAGAATGTTACACAGGTATTACT 600
DO	gagaaatctactacttaccaaatggaagaataacaagcagtaagatgtlccagcaagttact 609
OY	GGTACAAATGTGTGAAGAGAGTGTAGGGAGCAGCTGTATACCAAAAAGATGCACATTC 660
DO	ggtacaaatgtgtgaagaagatgtcagggagcagctgtttacaccaaagaatgtcacattc 669
OY	AATAAAGTAAACGAGACTCTATGAAAAAGATGTGAGTTTGTTCTCCATTCGCCGACAGC 720
DO	aataaagttaacagagactctatgaaaaagatgttgagttgttcttccaatccgcgcagaag 729
OY	CAGAGGCGTTCTATATGTGTTSCAACAACTGTGATTCTATAGTTCAATTCCTGTACAGA 780
DO	cagagagcttctataatgttttgacaaacatgttgattctcatatglttgaattctgtacaaga 789
OY	CAAAACCAACAAAGAAAGCTCCAAACAGCAAAATCAAAAATGCAATCTCCGAAGCAC 840
DO	caaaacccaacaaagaagctcacaacagcaaaaatcaaaaatgtcaatctccgaagcacaa 849

841 TGGGAAGTATCCGATTTCTGAGCACTTTAAGAAAACCACTCTATGACAAACAGCCCA 900
850 Tgggaagtgatccgctgattctctgaggaacttaagaaaacacccctatctgcaacacagcca 909
901 CCAATCCCACTTCTCATCTGCTGCAAGATTTGACAAAGATTTGTTAGTCTTGCAC 960
910 ccaatcccaactctctctgctgagatlgacaaagatgltglttaagtccttgagc 969
961 AAATGTGAGCATGGCGCTGTTAACCGCTCAATCGCTGAATCAACAGCCGACGCT 1020
970 aaatctggaagacagcagctcagctcagccccaacgacccaagcagcagcct 1029
1021 TTCTGTCTGACAGACAGTTGAGCTGGGGTCTCTGGTTGGAGTGTGACATTTGACAGTCT 1080
1030 ttctgtctgacagacagttgagctggggtctctgggttgagtgatcttgacagctgct 1089
1081 GCCCATGTACAAAGTGAAGTCTATACAGATTAACAGTGGCAGTGAACAGCACTCGCC 1140
1090 gcccatgtacaaagtgaagctcatacagatlaaacagtgagcagtgacagagcacactcgcc 1149
1141 AAAAGATTACCTGACAGCACTTACAGAGGAGCTCCATCTGACAGCGCTTGTGATCGGCA 1200
1150 aaaagattacctgacagcagctcagcagagagagcagctcagcagcgagctcgatcgca 1209
1201 TTTACTGTATTAGGAGCAAAATATCCAACTGATGATCTGAATTTGCTGCTGACGGAT 1260
1210 tttaactgtattaggaagaataataccaactgtatcgaatctgtgtgtctgacgagat 1269
1261 GGGGAAGCAACACATTAAGTGGTGTCTTAAAGAGGTCAAAAGTGGTGGTCCATCATC 1320
1270 ggggaagcaaacacataagtggtgtctttaaagaggtcaaaagtggtgcccatacc 1329
1321 CACACAGTCGCTTTGGGGCTCTGACGCTCAAGACATGAGAGCTGTCCAAATGACA 1380
1330 cacacagtcgctttggggctctgacgctcaagacatgagagctgtccaaatgaca 1389
1381 GGAGGTTTACAGACATATGCTTCAGATCAAGTTCAAGACATGGCTCATTTGATGCTTT 1440
1390 ggaggtttacagacatattgcttcagatcaagttcagaacaatgagctcatcttgatcttt 1449
1441 GGGGCCCTTTCTCATGAGAAATGAGCTGTCTCAAGCGCTCCATCCAGCTTAGAGTAAG 1500
1450 ggggccctttctcatgagaaatgagctgtctcaagcgctccatccagcttagagtaag 1509
1501 GGATTAACTCCAGACAGCAAGCTGATGATGAGACAGTATGATGAGACAGCCGCTG 1560
1510 ggattaaactccagacagcaagctgagatgagacagtgatcgtgagacagcagcgtg 1569
1561 GGAAAGCACTTTGTTCTTATCACCTGACACAGCAAGCTCCCAATCTCTCTGG 1620
1570 ggaagacacttgttcttatacactgacacagcagcagcctcccaatctctcttg 1629
1621 GATCCCATGTGACAGACAGCAAGCTGGCTTTGATGAGACAAAACCAAAATGGCTTAC 1680
1630 gatcccatgtgacagacagcaagctggctttgatgagacaaaaacaaaatggcttac 1689
1681 CTCCAATCCCGAGCATTTGCTTAAGTTGGCACTTGGAAATACAGTCTGACCAAGCTCA 1740
1690 ctccaatcccgagcatttgcctaagttggcacttggaataacagctcgaaacagcctca 1749
1741 CAAACCTTGACCTGATGCTGACAGTCCGTCGCTGCAATGCTACCTGCTCCATTTACA 1800
1750 caaaccttgacctgactgacagctcgcgtgctgctcaatgctacactgctccatata 1809
1801 GTGACTTCCAAAAGCAACAGACACAGCAAAATTCGCCAGCTCTGGTAGTTATGCA 1860
1810 gtgacttccaaaagcaacagacacagcaaatcccgagcctctgtagttatgca 1869
1861 AATATTCGCAAGAGCGCTCCCAATTCCTCAGGCGCAGTGTACAGCCCTGATTAATCA 1920
1870 aatattcgcaagagcgctcccaattctcagggccagtgctacagcctgattgata 1929

1921 GTGAATGAAAAACAGTTACTGTAAGTACTGATTAATGAGCAGAGTCTGATCTACT 1980
1930 gtgaatgaaaaacagttactgtaagtactgataatgagcagagtgatctgact 1989
1981 AAGGATGACCGTGTCTACTCAAGGATTTTACAACTTATGACACGAATGTTGATACAT 2040
1990 aaggatgaccggtgtctactcaaggattttacaacttataagacgaatgtagatacag 2049
2041 GTAAAAGTCCGGCTCTGGAGAGAGTTAAGCAGCAGCAGGAGTGTATACCCAGCAG 2100
2050 gtaaaagtccggctctggagagagtttaagcagcagcagagagatgatacccaagc 2109
2101 AGTGAAGCACTGTACATACCTGGCTGATGATGAAATGATGAATACATGAATCCACCA 2160
2110 agtgaagcactgtacatacctgctgagatgagatgataatacaatgataccacca 2169
2161 ACAGCTGAAATTAATTAAGATATGTTCAACCAAGCAAGTGTGTTTACGCAAGAACATCC 2220
2170 agacctgaaatlaataaagatgatttcaacacaagcaagtglttctcagcagaacatcc 2229
2221 TCGGAGGCTCATTTGTGCTCTGATGTCCTCAATGCTCCCATACCTGATCTCTTCCA 2280
2230 tcggaggctcatttgtgctctctgattgtcccaaatgctccatctgattctccca 2289
2241 CCTGGCCAAATCACCGACCTGAGGCGGAAATTCACGGGGGAGCTCATTAATCTGACT 2340
2250 cctggccaaatcacccgacctgagggcggaattccacgggggagctcatctactgact 2349
2341 TGGACAGCTCCCTGGGGATGATATGACCATGGAACAGCTCACAAGTATATCTTGAATA 2400
2350 tggacagctccctggggatgatattgaccatggaacagctcacaaagtattcttgaata 2409
2401 AGTACAGATTTCTTGATCTGAGACGAAAGTCAATGAATCTCTCAAGTGAATGATCTACT 2460
2410 agtacagatcttctgattctcagagacaaagttcaatgaatctctcaagtgatactact 2469
2451 GCTCTATCCCAAGAAAGCAACTCTGAGAAATCTTTTGTAAACCAAGAAACATTT 2520
2470 gctctatcccaagaaagcaactctgaggaagcttcttctttaaaccagaaacatt 2529
2511 ACTTTGAAAATGCGACAGATCTTTTCAATGCTATTCAGGCTGTTGATTAAGTGCATCTG 2580
2530 actttgaaaatgcgacagatcttttcaatgctatcagcgtgtgataagtgatctg 2589
2581 AAATCAGAAATTAATCAACTTTCAGAGATCTTTGTTTAAATTCCTCAAGCTCCGCCA 2640
2590 aaatcagaaatlaatacaacttgcagagatatttgttatactccccaagactccgcca 2649
2641 GAGACACTAGTCTGATGAAACGTCGTCTGCTTCTCTAATATTCATATCAACAGCACC 2700
2650 gagacactagctcgtatgaacgctgctctgtcttcttaataataatacaacagcacc 2709
2701 ATTCTGGCATTCACATTTTAAATTAATGATGGAAGTGAATGAGAACTGACGCTGTCA 2760
2710 attctggcatctcacattttaaataatgtagaagtgagatgagaaactgacgtgtca 2769
2751 ATAGCCTAGGCTGAAATTTTGTGCAATTAATTAATTAATATTCATCTTCACTT 2812
2770 atagcctaggctgaaattttgtcagataataataataataatcaatccctt 2821

RESULT 4
AAH46124
ID AAH46124 standard; cDNA; 2825 BP.
XX AAH46124;
AC
XX
XX
XX 11-SEP-2001 (first entry)
XX
XX
XX human CLCA1 cDNA, SPO ID NO:26.
XX
XX
XX human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
XX expression inhibition; antisense therapy; gene therapy;

OY	1443	GGCCCTTTTCATCAGAAATGGAGCTGTCCTCAGGCGCTCCATCCAGCTTGAGATGAAGAAG	1502
Db	1441	ggcccttcacacaggaatggagctgtctctccagcgctccacccagcttggaagataaagg	1500
OY	1503	ATTAACTCTCCAGAACAGCCAGTGGATGAATGGCACAGTATCGTGGACAGCAGCCGTGGG	1562
Db	1501	attaacctctccagaaacagccagtgatgaatggcacagtgatctcgtggacagacagctggg	1560
OY	1563	AAAGGACACTTTGTTTCTTATCACCTTGACACAGCAGCTCCCAATCTCTTGTGGGA	1622
Db	1561	aaaggaacacttgttctctatcaacttggacaacagcagctccccaatctctctcggga	1620
OY	1623	TCCCGTGTGACAGAGCAAGAGGTGGCTTTTGATGGGACAAAAAACACAAAATGGCTTACTT	1682
Db	1621	tcccggttggacagagcaaggtggccttttgatgggacaaaaacacaaaatggcttacct	1680
OY	1683	CCAAAATCCAGGACTTGTCTAAGTTGGACATTGGAAATACAGTCTGACAGCAAGCTCACA	1742
Db	1681	ccaaatcccgagcatgtctaagtgtgaccttggaaatacagtlcttgcgaagagtcaca	1740
OY	1743	AACCTTGACCTGACTGTACGTCCGTGGCTGCATGCTACCCGTGCTCCATTTACGT	1802
Db	1741	aaccttgcaccttgactgtcacgtcccgctgctcgaatgtcaccttgcctccaatcagtl	1800
OY	1803	GACTTCCAAAACGAAACAAAGCACACACAAATTTCCACAGCCCTCGTGTGTTTATGAAA	1862
Db	1801	gacttccaaaacgaaacagaagacacacgaatctcccaagccctcgtgtgaatttagcaaa	1860
OY	1863	TATTGCCAAGAGACTCCCAATTTCAGAGGCGACTGTACAGCCGTGATGAATACGT	1922
Db	1861	tattgccaaagagcctccccaattctcagggcaggtgcacagccctgtatgaatcagtl	1920
OY	1923	GATTTGGAAAAACATTTACTTTGGAACTACTTGGAATATGACAGAGTGTCTGATGCTATTA	1982
Db	1921	gaatttggaaaaaacatttactttggaaactacttggaaatgatgacaggtgtctgactactaa	1980
OY	1983	GGATGACGCTGTACACCAAGGATTTTTCACAACTTTATGACAGGATGGATGATACAGCT	2042
Db	1981	ggaatgacgctgtacaccaaggtatttccaactatgacaggaatggatgatacagtgtl	2040
OY	2043	AAAAGTGGGGCTCTGGGAGAGATTAAAGCAGCAGCAGCAGAGATACCCACAGCAAG	2102
Db	2041	aaaagtggggctcttgggagagattaaagcagcagcagcagcagagatgatacccaagcagag	2100
OY	2103	TGGAGCACTGTACATACCTGGCTGGATTGGAATGATGAATATCAATGGATCCACCAAG	2162
Db	2101	tggagcactgtacataccttggcttggattggaatgatgaatatcaatggatcccaag	2160
OY	2163	ACCTGAAATTTAATGAAGATATGTTCAACCAACCAAGAGTGTTCACAGACAATCCTCT	2222
Db	2161	acctgaaatttaatgaagatatgttcaaccaaccaagagtggttccaagcagaatctctc	2220
OY	2223	GGGAGGCTCATTTTGTGCTTCTGATGTCCCAATAGTCCCAATCCATGATCTCTTCCAC	2282
Db	2221	gggagggctcatTTTgtgcttctgatatgtcccaatagttcccaatcactgtatctcttccacc	2280
OY	2283	TGGCCAAATCACGAGCTGAAGCGGAAATTCACGGGGCAGTCTCATTAATCTGACTTG	2342
Db	2281	tggccaatcacagactgaagcggaaattcacggggcagctctcatattactgtacttg	2340
OY	2343	GACAGCTCTCGGGATGATTAATGACCATGGAACAGTCAAGTATTAATCATTTGAAATAG	2402
Db	2341	gacagctctcggggaatgattatgacatlgaaacagctcacaagatatcatctcgtatag	2400
OY	2403	TACAAAGTATTTATGATCTCAGAGACAAAGTTCAATGAAATCTTCAAGGAATACACTGC	2462
Db	2401	tacaagtattcttctgattctcagagacaagtccaatgatactcttcaagtgatactactgc	2460
OY	2463	TCTCATCCCAAAGGACCAACTCTGAGGAAGTCTTTTGTTTAAACGAGAAACATTAC	2522
Db	2461	tctcatcccaaaaggaagccaactctgaggaagtccttltgtttaaccagaaaaaatctac	2520
OY	2523	TTTTTGAATAATGGCACAGATCTTTTTCATTTGCTATTCAGGCTGTTGATTAAGTGATCTGAA	2582

Dp	2188	aaatattcgcgaagagagccctccccaattctccaggcgagtgfcacagccctgattgaa	2247
Qy	1921	gtgatattgaaaaaacactttacccttggaaactacttggatattggacagcagctcgatgact	1980
Dp	2248	gttaaatgagaaaaacagcttaaccttggaaactactctgataatggagaggtgcctgatt	2307
Qy	1981	aaggaagacgggtgtctactcaaggaattttcacaactttagatggacagaaatgtagat	2040
Dp	2308	aaggaatgagcgggtgtctactcaaggaattttcacaactttagatggacagaaatgtagat	2367
Qy	2041	gtaaaaatgctgggctctgsgagagattaaacgacgccagacgagatgattaccacag	2100
Dp	2368	gtaaaagtgcggctctcggagaggttaacacagccaaagagagtgatccccagag	2427
Qy	2101	agtgagagcactgtacactgacctggctggattggaaatggaatgaaatggaatgcacca	2160
Dp	2428	agtgagagcactgtacactgacctggctggattggaaatggaatgaaatggaatgcacca	2487
Qy	2161	agacctgaaattatgaatgaatgttttaacacaaagcgaatgtgtttcagcagaaatcc	2220
Dp	2488	agacctgaaattatgaatgaatgttttaacacaaagcgaatgtgtttcagcagaaatcc	2547
Qy	2221	tcgggaagcctcattttgtggctgtctgaaatggcccaaaatggcccatgactgtgtccca	2280
Dp	2548	tcgggaagcctcattttgtggctgtctgaaatggcccaaaatggcccatgactgtgtccca	2607
Qy	2281	ccttgcccaaatatccagccctgaaagaggaattcaacggggagctgatttaattgact	2340
Dp	2608	ccttgcccaaatatccagccctgaaagaggaattcaacggggagctgatttaattgact	2667
Qy	2341	tggacacctctctgggagattgatttgaccatggaaacgctcacaaatgattatcttgata	2400
Dp	2668	tggacacctctctgggagattgatttgaccatggaaacgctcacaaatgattatcttgata	2727
Qy	2401	agtcacagttttcttgatgtcagacagacagtttcacgaatgcttctcaagtgaatctact	2460
Dp	2728	agtcacagttttcttgatgtcagacagacagtttcacgaatgcttctcaagtgaatctact	2787
Qy	2461	gctctcaccctcccaaaagagcccaactctgaggaagctcttctttaaaccagaaacatt	2520
Dp	2788	gctctcaccctcccaaaagagcccaactctgaggaagctcttctttaaaccagaaacatt	2847
Qy	2521	acctttaaaaatggcacaagatcttttcatttgcatcagcgttgatgaagtcgactcg	2580
Dp	2848	acctttaaaaatggcacaagatcttttcatttgcatcagcgttgatgaagtcgactcg	2907
Qy	2581	aaatcagaatattccaacatttggacagagatgcttcttattcttcacagactccgcc	2640
Dp	2908	aaatcagaatattccaacatttggacagagatgcttcttattcttcacagactccgcc	2967
Qy	2641	cagacacctgaattctgattgaaagctcgcctcttgcctaaatattcaatacaacgac	2700
Dp	2968	cagacacctgaattctgattgaaagctcgcctcttgcctaaatattcaatacaacgac	3027
Qy	2701	atttcctggcatttcacatt	2760
Dp	3028	atttcctggcatttcacatt	3087
Qy	2761	atttcctggcatttcacatt	2812
Dp	3088	atttcctggcatttcacatt	3139
RESULT 6			
AAH33285			
ID AAH33285 standard; cDNA: 2867 BP.			
XX AAH33285:			
XX AC			
XX DT 03-SEP-2001 (first entry)			
DE Human colon cancer antigen encoding cDNA SEQ ID NO:341.			

Query Match	99.5%	Score 2797.8	DB 22	Length 2867
Best Local Similarity	99.9%	Pred. No. 0		
Matches 2810	Conservative 0	Mismatches 2	Indels 1	Gaps 1
1	GAATTCACACGGGANGATGTCACAGCAATGGGGCCATTAAAGTCTGTGTCATCTGANT	60		
2	gaatcacacgggagagatgacagcaaaaggccattaaagatcgtgtcatcttgatt	73		
3	CTTCAACCTCTCTGAAGGGGCGCTGATGTAATTCATCTCATCTTCAGCTGAACAACATGGCTAT	120		
4	cttcacacctctgaagaggggcctgtgtaattcatctatctcagctgtaacaacaatgctat	133		
5	GAAGGATTTGTGTTGCATTCGACCCCAATGTGCGAGAGATGAACACTCATTCACAA	180		
6	gaaggattgtcgttgcaatcgaccacatgtgcagaagatgaacacatctaccacaa	193		
7	ATTAAGGACATGTGTGACCCAGCATCTCTGTATCTGTTGAGGCTCAGAGAAAGCATTT	240		
8	ataaaggaatggtgacccagagcatctcgtatctgttgaacctacaggaagagatt	253		
9	TATTTCAAAAAGTGTCCATTTTGATCTCTGAAACATGGAAGCAACAAGGCTGACTTGTG	300		
10	tattcaaaaagtgtgcattttgattcttcgtgaacaactggaagcaaaagctgactatgtg	313		

301 AGACCAAACTTGAGACTACCAAAATGCTGATGTTGCTTCTGATCTACTCTCCA 360
314 aagacaaacttgaagaccctacaaatgctgagttctgagttcgtactctcca 373
361 GGTAATGATGAACCTTACACTGAGAGATGGCACTGGAGAGAGAGGATGATC 420
374 ggttaatgataaaccctacacacagagacagtgaggaactgaggaagagtgaaagatc 433
421 CACCTCACTCTGATTTTCATTTGAGAGAAAAAAGTTAGCTGAATVATGACCAAGATAG 480
434 caactcaactcgtattcattctgaggaagaaatgactggaatattgagcaacaaagtag 493
481 GCATTTGTCATGATGAGGCTGCTACTGATGAGAGAGATTTGACAGATCAATTAATGAT 540
494 gaattgtcactatgagtgaggtcactcactagagtgaggtatltgaagagatcaaatgat 553
541 GGAATTTCTACTTATTCATGAGAGAGATGAGATGAGATGATGATGATGATGAT 600
554 gagaactcactcactcactcactcactcactcactcactcactcactcactcactcact 613
601 GGTACAAATGATGATGAAG 660
614 ggtacaatgtagtaagaagagtagcagagagagagagagagagagagagagagagagag 673
661 AATAAGTACAGAGACTCTATGAAAAAGAGTGTGAGTTGTTCTCCAAATCCGCGCAGAG 720
674 aataaagttacagagactcactcactcactcactcactcactcactcactcactcactc 733
721 GAGAAAGCTTCTATATGTTTGACAAAGATGTTGATGATGATGATGATGATGATGAT 780
734 gagaagagctcactcactcactcactcactcactcactcactcactcactcactcactc 793
781 CAAAACACACAG 840
794 caaaacacacaaagaagagctcactcactcactcactcactcactcactcactcactc 853
841 TGGAGAGTATCCGATGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
854 tgggaagtagatccgtagatctcactcactcactcactcactcactcactcactcactc 913
901 CCAAAATCCACCTTCTCATTTGCTGAGATGTTGAGAGAGAGAGAGAGAGAGAGAGAG 960
914 ccaaatcccaactcctcactcactcactcactcactcactcactcactcactcactcact 973
961 AAATCTGAG 1020
974 aaatctgagaagatgag 1033
1021 TTTCCTGCTGACAGAGTGAAGTGGGGTCTGGGTTGGGATGGTGAATTTGACAGTCT 1080
1034 ttcctgtgacagacaggttgaagctggggtctggttgggagtgatgacattgacagtgct 1093
1081 GCCCATGTACAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1139
1094 gcccatgtacaaagtgaaactcactcactcactcactcactcactcactcactcactcact 1153
1140 CAAAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1199
1154 caaaagattacactcagcagctcagagagagagagagagagagagagagagagagagag 1213
1200 ATTTCTGATGATGAG 1259
1214 atttctgattgattgagaagaaatcactcactcactcactcactcactcactcactcactc 1273
1260 TGGGAG 1319
1274 tgggag 1333
1320 CCACAGAGTGGCTTTGGGGCTCTGACAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1379
1334 ccacagagtgctttggggcctctgacagctcaagagagagagagagagagagagagagag 1393
1380 AGGAGGTTTACAGACATATGCTTCAAGATCAAGTTCGAAACATGGCTCATTTGATCTTT 1439

1394 aggaagttacagacatattcttcagatcaagtagcagaacaaagcctcattgagcttt 1453
1440 TGGGGCTTTTCACTAGGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1499
1454 tggggcctttcactcaggaatgagagctgtctcctcagcgtccatccatccagttgagtaa 1513
1500 GGGATTAAACCTTCAG 1559
1514 gggattaaaccttcag 1573
1530 GGGAAAGGACACTTTGTTTCTTATGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1619
1524 gggaaaggaacttgttcttactcactcactcactcactcactcactcactcactcactcact 1633
1620 GGTATCCAGTGGAG 1679
1634 ggtatccagtgag 1693
1680 CCTCCAAATCCAG 1739
1694 cctccaaatcccaagagatgtag 1753
1740 ACAACCTTGACCTTGAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1799
1754 acaaccttgacctgactgtcactcactcactcactcactcactcactcactcactcactcact 1813
1800 AGTGAAGTCCAAAG 1859
1814 agtgaagttccaaag 1873
1860 AAATATTTGGCAG 1919
1874 aaatatttgccaag 1933
1920 AGTGAATGAG 1979
1934 agtgaatgag 1993
1980 TAAGATGAG 2039
1994 taagatgag 2053
2040 TGTAAAGTGGCGGCTCTGGAG 2099
2054 tgtaaagtggcggtcttgag 2113
2100 GAGTGAAGCACTGTACATACCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2159
2114 gagtgaagcactgtacatacctgtgactgagatgagatgagatgagatgagatgagatgagat 2173
2150 AAGACTGAATTAATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2219
2164 aagactgaattaatataagaatgagatgagatgagatgagatgagatgagatgagatgagat 2233
2220 CTGCGAGAGCTCATTTTGGGCTTCTGATGCTCCAAATGCTCCATACCTGATCTCTTCCC 2279
2234 ctgcgagagctcatTTTGGGCTTCTGATGCTCCAAATGCTCCATACCTGATCTCTTCCC 2293
2290 ACCTGGCCAAATACAG 2339
2294 acctggccaaatcacag 2353
2340 TTGGAGAGCTCTGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2399
2354 ttggagagctctgggagagatgagatgagatgagatgagatgagatgagatgagatgagatgag 2413
2400 AAGTACAAATTAATTTGATCTGAG 2459
2414 aagtacaagtatattctgtatctcagagagagagagagagagagagagagagagagagagag 2473
2460 TGCTTCATCCCAAG 2519
TGCTTCATCCCAAG

Db 2474 tgcctcaccacaagaagcacaactcgtgagaagctcttctgtttaaccagaaccat 2533
QY 2520 TACTTTTGAATAAGGACAGATCTTTTCATGTCATTCAGGCTGTGTATGAGTCGATCT 2579
Db 2534 tactttgaaatggaagacagatcttctcattcattcagctcgttgcataagtcgactc 2593
QY 2580 GAAATCAGAAATATTCACATTCGACGAGATATCTTTTATTCCTCCACAGACTCCGCC 2639
Db 2594 gaattcagaataatccacacttgcagctatcttctgttattctccacagactccgcc 2653
QY 2640 AGAGACACTGTCTGTATGAAAGCTCTGCTCTTGTCTTAATATTCATTCACAGACAC 2699
Db 2654 agagacactcgtctcgtatgaaagctcgtccttgccttaataatcattcacaagcac 2713
QY 2700 CATTCCTGGCATTCACATTTTAAATAATATGATGAGAGAGAGACTGCAGCTGTC 2759
Db 2714 catctcgtgacatcattcattcaaatatcgtgaaagtgaaagagacgtgacgtc 2773
QY 2760 AATAGCCAGGCGTGAATTTTTCAGATTAATAATTAATTAATTCATTCCTT 2812
Db 2774 aatagcctaggcgtgacttcttgcataataataataatcattcactc 2826

RESULT 7
AAF81927 ID AAF81927 standard; cDNA; 2745 BP.
XX
AC AAF81927;
DF 13-JUN-2001 (first entry)
XX
DE Human ICACC-1 nucleotide sequence.
XX
KW ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
KW interleukin 9 induced calcium activated chloride channel; IL-9;
KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
KW anti-inflammatory; immunomodulatory; cystic fibrosis;
KW inflammatory bowel disease; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH key Location/Qualifiers
FT 1..2745
FT /*tag= a
FT /product= "ICACC-1"
FT /note= "IL-9 induced calcium activated chloride channel"
XX
PN M09944620-A1.
XX
PD 10-SEP-1999.
XX
PE 03-MAR-1999; 99WO-US04703.
XX
PR 03-MAR-1998; 98US-0076815.
XX
PA (MAGA-) MAGALININ PHARM INC.
XX
PI Holroyd KJ, Levitt RC, Maloy WL, Louhed J, McLane M;
PI Nicolides NC, Zhou Y, Dong Q;
XX
DR WPI: 1999-550979/46.
DR P-PSDB: AAB74824.
XX
PT New nucleic acid encoding calcium activated chloride channel, used to
PT identify, e.g. specific modulators for treating atopic allergy -
XX
PS Claim 1; Fig 4B; 75pp; English.
XX
CC The present sequence encodes the human interleukin 9 (IL-9) induced
CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
CC have anti-allergic, anti-asthmatic, anti-inflammatory and
CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
CC used to alleviate asthma (or more generally atopic allergy), while those

CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma
CC (increased levels) or IBD (reduced levels), also for monitoring
CC treatment of these conditions. The ICACC proteins can be used:
CC (1) to raise specific antibodies (Ab), useful: (a) as immunoassay
CC reagents, and (b) as therapeutic (A); (11) as fragments or mutants, as
CC (A) to disrupt bonding between ICACC and its ligands; and (111) to
CC identify modulators and binding partners. ICACC polynucleotides can be
CC used to generate transgenic animals or recombinant cells, used to screen
CC for antagonists, also as a source of therapeutic antisense agents or
CC diagnostic probes (for quantifying mRNA expression, e.g. for
CC identification of modulators).
XX
SQ Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;

Query Match 97.5%; Score 2743.4; DB 20; Length 2745;
Best local similarity 100.0%; Pred. No. 0;
Matches 2744; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 ATGGGCGCATTTAAGATCTGCTGATCTTGAATTCATTCACCTTCTGAGGGCCCTG 84
Db 1 atgggcccattaaagagctcgtgtcattcctcaccctcagaaggccctg 60
QY 35 AGTAATTCATTCATTCACCTGACACAAATGCTATGAGGCAATGCTGCAATCGAC 144
Db 31 agtaattcattcattcagctgaaacacatgctatgaaagcatgtgttgcacatgac 120
QY 115 CCCAATGTCGCAAGAGATGAAGAACTATTCACAAATTAAGACATGTCACCGGCA 204
Db 121 cccaatgtgccagaagaatgaacacatctcaacaaataaagacatgtgaccagaca 180
QY 205 TCTCTGATCTGTTGAAGCTACAGAAAGGATTTTTCAAAATGTTGCCATTGTG 264
Db 181 tctctgattcgtttgaaagctacagaaagacatcttctcaaaaagttgcacatttg 240
QY 235 ATTCTGGAACATGGAAGACAAAGGCTGATGAGACCAAAATTTGAGACCTTACAA 324
Db 241 attctgaaacatggaagaacaaagctgactatgtgagcaaaaacttgagacctaca 300
QY 335 AATGCTGATGTTGCTGCTGAGCTGATCTCTCCAGTAATGATGAACCTTACACTGAG 384
Db 331 aatgctgattgttctgtgtcgtgactcctccagtaatgataagacccctacactgag 360
QY 335 CAGATGGCAACTGTGAGAGAAAGGTAAGAGATCCACTCATCTCGATTTCATGCA 444
Db 331 cagatggcaactgtgagagaaaggtgaaagatccaccctcactcctgatttcacga 420
QY 445 GGAAAAAGTTAGCTGAATATGACACCAAGGTAGGCAATTTGTCATGAGTGGCTCAT 504
Db 421 ggaanaaagttagctgaataatgacacaaagtgaggaattgtccatgagtgctcat 480
QY 505 CTACATGGGAGATATTTGACGAGTACATTAATGATGAATTTCTACTTCCAAATGGA 564
Db 481 ctacatgggagatatttgacgagtaataatgataatgtgaattctactctccaatgga 540
QY 535 AGAATACAGAGATGATGATGTTCCAGCAGTATTACTGTACAAATGTGTAAGAAGTGT 624
Db 511 agaatacaagcagtaagaatgttcagcaggtattactgttacaatgttaagaagtggt 600
QY 635 CAGGAGGCGAGCTGTACACCAAAAAGATGACATTCATTAAGTAACAGACTGTATGAA 684
Db 621 caggagggcagctgttacacaaaagatgacattcataaagttacagagctcatgaa 660
QY 635 AAGAGATGATGATGTTGTTCTCCAATCCCGCAGACAGAGGAAGGCTTCATATATGTTT 744
Db 631 aagagatgattgttctccaaatcccgccagagaggaaggtctctataatgtttgca 720
QY 745 CAACATGTTGATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 804
Db 721 caacatgttattctatgttgaattcgttacagacaaaacacacaaagaagctcca 780

805 AACAGCAAAATCAAAAATGCAATCTCCAGACATGGAGAACTGATCCGTATCTGAG 864
|||||
781 aacagcaaaatcaaaaatgcaatctccgaagacaatgggaagtgtccgtgtcttctg 840
865 GACTTAAACAAAACCTCTCTATGACACACAGCCCAATCTCCACTTCTATGCTG 924
|||||
841 gactttaaagaaacccctctatgacaacacagcccaatccctctctatctgtc 900
925 CAGATTGGCAAAAGATTGTTGTTAGTCTGACAAATCTGGAAGCATGGGACTGGT 984
|||||
901 cagattggcaaaagattgtgtttagtcttgacaacatctggaaagcatggcagctgt 960
985 AACCGCTCAATCGACTGATCAAGCAGGCGACCTTTCTCTGCTGACAGACTTGAAGT 1044
|||||
961 aaccgctcaatcgactgactgataagcagccagcttctctgtctgagaaagtgtg 1020
1045 GGGTCTGGGTTGGGATGGTGTGACATTTGACAGTCTGCCATATCAAAAGTGAATCAT 1104
|||||
1021 gggctctgggttgggattgggtgacatttgacagtgctgcccacgttacaagtgacata 1080
1105 CAGATTAACAGTGGAGTGCAGAGACACACTGCCAAAAGATTACCTGCAGAGCTTCA 1164
|||||
1081 cagataaacagtgagcagtgacagggacaacatcgcaaaagattacctgcagacttca 1140
1165 GGAGGACCTCCATCTGACAGCGGGCTTGCATCGCATTTACTGTGATTAGCAAAATAT 1224
|||||
1141 ggaaggagcttcaatctgcagcgggtctgactgcaattactgtgtttaggaagaaat 1200
1225 CCACCTGATGATCTGAATTTGCTGCTGACGATGGGAGAACACACTATTAAGTGGG 1284
|||||
1201 ccaactgatagtatctgaatctgtcgtctgacgagatgggaaagacaactaaatgtg 1260
1285 TGCTTAAAGAGGTCAAAAGATGGTGCATATCATCAGCATGCTGCTTGGGGCCCTC 1344
|||||
1261 tgcctttaaagaggtcaaaaagtggtgcataccacacagtcgttggggccctc 1320
1345 GCAGCTCAAGAACTAGAGAGAGCTGTCCAAAATAGACAGAGGATTAGACATATGCTCA 1404
|||||
1321 gcaagctcaagaactagagagagctgtccaaaatgacagaggtttacagacatagtctca 1380
1405 GATCAAGTTCAAGACAATGGCTCATTTGATGCTTTGGGGCCCTTTCATACAGAAATGGA 1464
|||||
1381 gatcaagttcaagacaatgagctcatatgacttcttggggcccttcatcaagaaatgga 1440
1465 GGTGTCTCAAGGCTCCATCCAGCTTGAAGATTAAGGATTAACTCCAGAACAGCCAG 1524
|||||
1441 gctgtctcagagctccatccagcttggagtaagagatlaacacctccgaagcagcag 1500
1525 TGGATGAATGGCAGCTGATCTGTGACAGCAGCCGTGGAAAAGACACTTTGTTCTTATC 1584
|||||
1501 tggatgaatggcagctgattgtgacagcagccgtggaaaagacactttgttcttctatc 1560
1585 ACCTGACAAACGAGCTCCCAAAATCTCTGTGGATCCCACTGACAGCAAGCAAGGT 1644
|||||
1561 acctgacaacagagcctcccaaatctctctgtggtatccagtgagcagaagcaaggt 1620
1645 GCGTTTGTAGTGACAAAAACCAAAAATGGCTTACCTCAAAATCCAGAGCATTTGTAAG 1704
|||||
1621 ggccttgtgtgtgacaaaaaacacaaaatgtgcttaacctcaaatccagcattgtctaa 1680
1705 GTTGGCACTTGAATACAGTCTGCAAGAGCTCACAAAACCTTGACCTGACTGTACG 1764
|||||
1681 gttggcacttggaaatatacagctctgcaagcaagcttcaaaaaccttgacctgtctacg 1740
1765 TCCCGTGGCGTCAATGTACCTCTGCTCCAAATTAAGTACTGACTTCCAAAAGCAAGGAC 1824
|||||
1741 tcccggtggcgctcaatgtactacccgtctccaatcagtgacttccaaaagcaagacgac 1800
1825 ACCAGCAAAATCCCAAGCCCTGTGTAGTTATGCAAAATTTGGCCAAAGAGCCCTCCCA 1884
|||||
1801 accagcaaaatccccaagccctgtgtagtgtttagtcaaatatctgcgaagagcctcccca 1860
1885 ATTCTCAGGGCCAGTGTCAACGCCCTGATTGAATGATGAATGCAAAACAGTTACCTTG 1944

1861 attcagggccagctgtcaacagccctgattgaatcagtgatgtaaaaaagttactctg 1920
1945 GAACACTGATTAATGAGACAGAGTGTGATGATGATGATGATGATGATGATGATGATG 2004
|||||
1921 gaactactgtataatgagagagtgctgtgactactaagaatgagcgtgtctactcaag 1980
2005 TATTGAGAACTTATGACAGATGATGATGATGATGATGATGATGATGATGATGATG 2064
|||||
1981 tattcaaacacttatgacacgaatgltagatacagtgtaaaatgtgggtctcgggaag 2040
2065 GTTAACGACCAAGCAGAGAGTGTATACCCAGACAGAGTGAGGACACTGTACATACCTGGC 2124
|||||
2041 gttcaacgacagcagagagtgatgatacccaagcagagtgagcactgtacatacctgac 2100
2105 TGGATTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
|||||
2101 tggattgagaaatgataaatacaatggaatcccaagaccgtgaatataaagatgat 2160
2195 GTTCAACACAGCAAGAGTGTGTTTACAGCAGAACATCCCTCGGAGGCTCATTTGTGCTCT 2244
|||||
2161 gtccaacacagaagatgtgttccaagaaatccctcgggaagctcattgtgtgtctc 2220
2215 GATGTCCCAAAATCTCCATACCTGATCTCTTCCACCTGCGCAAAATCACCACTGAG 2304
|||||
2221 gatgtcccaaatgtctccatactgactctctccacctgtgccaatcacagacctgaag 2280
2305 GCGGAAATTCACGAGGAGCATCTCATTAATCTGATGAGACCTCTCTGGGGAGTATAT 2364
|||||
2281 gcggaatctacagggagagctccatataatctgactgtgacagctcctgaggatgat 2340
2365 GACCATGAGACACCTCAGATATATCATTCGATTAAGTAAAGTATCTTGAATCTGAGA 2424
|||||
2341 gaccatgagacacgctcaacagatataatcagataagatacaagatcttctgtactaga 2400
2425 GACAAGTTCAAGTAATCTCTCAAGTGAATAGTACTGCTCTCATCCCAAGAACCCAAC 2484
|||||
2401 gacaagttcaatgatactctcctaagtgatactactctctcctcccaaggaagccaac 2460
2485 TCTGAGGAAGTCTTTTGTGTTAAACCGAANAACATTACTTTAAATGGCACAGATCTT 2544
|||||
2461 tctgaggaagctcttctgtttaaaccagaataacttcttgaatgtgacagactct 2520
2545 TTGATTGCTATTCAGGCTGTTTGAATAGTGATGATGATGATGATGATGATGATGATG 2604
|||||
2521 ttaattgtctatcagagtggtgtgataagtgatctgaaatcagaataataccaatgtca 2580
2605 CGAGTATCTTTGTTTATTCCTCCACAGACTCCGCGACAGACACCTAGTCTGATGAAGC 2664
|||||
2581 cgaatctctgttattctctccacagactccgacagagacactagctcctgataagacg 2640
2665 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2724
|||||
2641 tctgctcctgttcccaataatcaatacaagaacacacttctcgtgacttccatcttaaa 2700
2725 ATTATGTGAAGTGAATGAGAGAACTGACAGTGTCAATTAAGCTAG 2769
|||||
2701 attatgtgaaagtgtataggaactcagctgtcaatagctctag 2745

RESULT: 8
AAH46102
ID AAH46102 standard; DNA; 2742 BP.
XX
AC AAH46102:
XX
DE 11-SEP-2001 (first entry)
XX
KW Human C1CAL coding sequence, SEQ ID NO:4.
KW Human C1CAL: goblet cell; mouse Gob-5 orthologue; drug screening;
KW expression inhibition; antisense therapy; gene therapy;
KW chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;

QY	1465	GCCTCTCTCAGGCGCTCACACAGCTTGAGAGTAAAGGGATTAACTCTCAGAACAGCCAG	1524
Db	1441	gctgtctcctaagcgctcccatccacgctctggagtaagysgttaaccctccgaagacgagcag	1500
QY	1525	TGAGTGAATGGCACAGTGAATGCTGGACACACCGTGGGAAAGACACTTTGTCTTATC	1584
Db	1501	tgyatgaaatgycacagatgcatctgycgacagacacgctggaaagacacttcttctatc	1560
QY	1585	ACCTGGACAACGACGCTCCCGCAAAATCCCTTCTGGAGATCCACAGTGGACACAGACAGCT	1644
Db	1561	acctggacaacgacgctccccaatcctctctggatcccaatggacagaaacgaagt	1620
QY	1645	GGCTTTTCTAGTGGACAAAACACCAAAATGGGCTCACTCCCAATCCAGGAGATTGCTAAG	1704
Db	1621	gaccttgtagtggacaacaaacacacaaatgagcttaacctccaaatcccaagcatgtctaa	1680
QY	1705	GTTGGCACTTGGAAATACAGTCTGCGAAGCAGCTACAAACCTTAGCCTGACTGTACAG	1764
Db	1681	gttgcaccttggaaatacagctctgcacaagaagctcacaaaccttgacctgactgcag	1740
QY	1765	TCCCGTCCGTCACATGCTACCCGCTCCCAATTACAGTACGTCTCCAAACGACAAAGAC	1824
Db	1741	tcccggtgcgtccaaatgctacacctgcctccaatlaagatgacttccaaagacaaagac	1800
QY	1825	ACCGACAAATTCGCCACAGCCCTCTGTAGATTATGCAAAATATTCGCAAGAGACCTCCCA	1884
Db	1801	accgacaattcccccagccctctgtgagttaatgcaaatatctgcgaagagccctcccca	1860
QY	1885	ATTCTCAGGGCGACAGTGTCAACGCCCTGATTGAATCAGTGAATGGAATAACAGTTACCTTG	1944
Db	1861	attctcaggccagtgctacaagccctgattgactgaatgaaagaaacagttactcttg	1920
QY	1945	GAATCTACTGGATTATGGAGCAGCTGCTGATGCGTAAAGATGACAGGAGTGTACGCAAG	2004
Db	1921	gaactactgtaataatggagcagtgctgattgactaaagaaagacagtgctactcaag	1960
QY	2005	TATTTCAACACTTATGACACAGAAATGGTAGATACAGTGAATAAGTGGGCTCTGGAGGA	2064
Db	1981	tatttcaactatagacaagaaatgtagatacagtgtaaaagtgcggctctcggagga	2040
QY	2065	GTTTAAACGACACAGAGGAGAGATACCCACAGAGAGGAGGACCTGTACATACCTGGC	2124
Db	2041	gttaacgagccacagagagatgataccacagcagaatggagcaactgtaactaactggc	2100
QY	2125	TGAGTTGAGATGATGAATAATACATGGAATCCACCAAGACCTGAAATTAATTAAGATGAT	2184
Db	2101	tgagttagaataatgaaatacaatgaaatccacaaagccctgaaatlaataagatgat	2160
QY	2185	GTTCAACACAAAGATGTGTTTCACACAAACATCTCGGGAGGCTCATTTGTGGCTTCT	2244
Db	2161	gttcaaacacaagaagaatgtgttctcagcagaacactcctcggaggtcatattgtgctct	2220
QY	2245	GATGTCCCAATGCTCCCAATACCTGATGCTCTTCCCACTGGGCAATACAGCACTGAG	2304
Db	2221	gatgtcccaatgtcccaataactgtctctcttcccaactcggcacaataccgacctgag	2280
QY	2305	CGCGAAATTCACGGGGGCACTCTCATTTAATCTGACTTGGACAGCTCTGGGAGATGATTAT	2364
Db	2281	cgcgaaatccaaggggcaatccataactcgaacttgacagcctccctggagatgatat	2340
QY	2365	GACCATGGAACAGCTCACAGTATATCATTTGGAATTAAGTACAAAGTATCTTGATCTCGA	2424
Db	2341	gaccatggaacagctcccaagatatcaatcgaataagaacaaagatattcttgcataccga	2400
QY	2425	GACAAAGTTCAATGAATCTCTTCAAGTGAATACTACTGCTCATGCCAAAGAACCCAAC	2484
Db	2401	gacaagttcaatgaaatctcttccaatgaaatactactgctcatccccaagaagcaac	2460
QY	2485	TCTGAGAGAACTCTTTTGTTTAAACACGAAACAACTACTTTTGAANAATGGCACAGATCTT	2544
Db	2461	tctgagaaagctcttctgtttaaaccagaacaaactacttggaaatggaacagactct	2520

QY	2555	ITCATTGGATTCACAGCCTGTGATAGTCGATCGAAATATCAAAATATCCAACTATGCA	2604
DP	2581	ttcatgtccattccaggctgctgtgataagcgtgcgactctgaataatcgaataatccacattgca	2580
QY	2635	CGAGTATCTCTTTGTTTATTTCTCCACAGACTCCGCCAGAGACACCTGATGCTGATGAACG	2664
DP	2641	tctgcgccctgtccchaataltcatatcaacagacacattccgtgcatlccacatttcaaa	2700
QY	2735	ATTATGTGGAGTGTGATAGAGAGAACTGACAGTGTCTCAATAGCC	2766
DP	2731	attatgtggaagtgtgataagtagagacactgcagctgtcataatgcc	2742
RESULT	9		
ID	AAH35019	standard; cDNA; 3109 BP.	
AC	AAH35019;		
LT	03-SEP-2001	(first entry)	
DE	Human colon cancer antigen encoding cDNA SEQ ID NO:2101.		
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection;		
KW	colorectal carcinoma; ss.		
OS	Homo sapiens.		
PN	WO200122920-A2.		
RD	05-APR-2001.		
XX	28-SEP-2000; 2000WO-US26524.		
XX	29-SEP-1999; 99US-0157137.		
RR	03-NOV-1999; 99US-0163280.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
FA	Ruben SM, Barash SC, Birse CE, Rosen CA;		
FI	Wp1; 2001-235357/24.		
DR	P.5SDB; AAG75614.		
XX	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,		
ET	useful for preventing, diagnosing and/or treating colorectal cancers -		
XX	Claim 1; Page 3587-3588; 9803pp; English.		
XX	AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon		
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where		
CC	the proteins are collectively known as colon cancer antigens. The colon		
CC	cancer antigens have cytostatic activity and can be used in gene		
CC	therapy and vaccine production. N and P may be used in the prevention,		
CC	diagnosis and treatment of diseases associated with inappropriate P		
CC	expression. For example, N and P may be used to treat disorders		
CC	associated with decreased expression by rectifying mutations or deletions		
CC	in a patient's genome that affect the activity of P by expressing		
CC	inactive proteins or to supplement the patient's own production of P.		
CC	Additionally, N may be used to produce the colon cancer-associated ps,		
CC	by inserting the nucleic acids into a host cell and culturing the cell		
CC	to express the proteins. N and P can be used in the prevention, diagnosis		
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204		
CC	and AAH77789 represent sequences used in the exemplification of the		
CC	present invention.		
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were		
CC	missing at time of publication, meaning no sequences are present for		
CC	SEQ ID NO:1027 to 1052, 7921 and 7922.		
XX			

Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match 93.2%; Score 2622.6; DB 22; Length 3109;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2623; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
QY 180 AATAAGGACATGTGATGACCCAGCATCTCTATCTGTTTGAAGGTACAGAAACGATT 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43 aatcgcagcaggtgaccasagcatctctglatctgttgaagctaaagaagcgatt 102
QY 240 TTATTTCAAAAATGTGGCCATTTTGTATTCCTGAACATGGAAGACAAAGCTGACTATGT 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 ttatltcaaaaatgttgccatttgatcttcgaaacalgyaagaagcgtgactatgct 162
QY 300 GAGACCAAACTTGAGACCTACAAAATGCGATGCTGTGTTGCTGATCTACCTCC 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 gagaccaaactlgyagactcaaaaatgctglatgctgtgctgactactctcc 222
QY 360 AGTAATGATGACCCCTACACTGAGCAGATGGCACTGTGGAGAGAGGTTGAAGGAT 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 agtaatgatagaacctactactgacagatggygcactgtgagagaaggtgaaagat 282
QY 420 CCACCTCCTCTCTGATTTGATTCAGAGAAAAAGTTAGCTGAATTTGACACCAAGTAG 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 ccacctcactcctgaltcaltgacagaaaaagtlagctgaatatgacacaaagtag 342
QY 480 GGCATTTTCCATGAGTGGGCTCATCTACGATGGGAGTATTTTGACGATCAATATGA 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 343 ggcatttgcatagaaggggtcactcagatggaggtatltggcaggtacaaataga 402
QY 540 TGAGAAATTTCTATTATCAATGAGAAATACAGACATGAAGTTTACAGCATTTAC 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 403 tgaagaattcacttatcctcaatgagaataataacagcagtaagatgtltagcaggtattac 462
QY 600 TGTGTAATTTAGTAAAGAGAGTGTGAGGAGGCGCTGTACACCAAAAGTGCACATT 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 463 tggtaaaaatgtaaaagaagtgcaaggagcagctgttacaacaaaatgacactt 522
QY 660 CAATTAAGTAAACAGACCTCTATGAAAAAGATGATGTTTTCATCCCGCCAGAC 719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 523 caataaagtwcagagactcctatgaaaaagatgtagtlttcttccaatcccgccagac 582
QY 720 GGAGAAAGCTTCTATATGTTTGAACAACATGTGATTCTATAGTTGAATTTCTGTACGA 779
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 583 ggagaagcttctataaagtlttgcaacaactgtgattctatagttgaattctgtacaga 642
QY 780 ACAAAACCAACAAGAAAGTCCAAACAAAGCAAAATCAAAATGCAATTCGCCAGCAC 839
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 643 acaaaacacacaaaagaagctccaacaaacaaatcaaaaatgcaatcttcgaaagcac 702
QY 840 ATGGGAAGTGAATCCGCTGATTTGAGGACTTTAAGAAAACCACTCCTATGACAAACACCC 899
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 703 atgggaagtgatcgtgacttctgaggaactttaaagaaacacactctatgacaacacagcc 762
QY 900 ACCAAATCCACCTTCTCATTTGCTGACAGATTGACAAAGAAATGTGTGTTTATGTCCTTGA 959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 763 accaaatccacacttctcatctgctgacagattgycacaaagaattggtttagtcttga 822
QY 960 CAATTTCTGGAAGCATGGGACTGGTAACCGGCTCAATGACTGAATCAAGACGGCAGCT 1019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 823 caaatctggaagcaltgagactgttaaccgcctcaatgactgaatcaagaagcagact 882
QY 1020 TTTCTGCTGAGACAGTTGAGCTGGGCTGCTGGGTTGGATGGATTTGACATGTC 1079
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 883 ttctcgtctgagacagtgtagctggyggtctctggttctgagatgtgaaattgaaagtcgc 942
QY 1080 TGCCCATGTACAAAGTAACTATACATACATAAAGTGGAGTGAACAGGACACACTGCGC 1139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 943 tgcacatgtacaaagtgaactcatacagataaacaagtgagcagtgagacacactgcg 1002
QY 1140 CAAAAGATTACTGACGAGCTTCAGGAGGAGCGTCCATCTGCACGGGCTTTGATCGGC 1199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1003 caaaagattactctgacagacttcagagagacgtccacatctgcacagcgagcttgcacgcg 1062
QY 1200 ATTTCGTGATTTGAGGAATAATATCCAACTGATGATGATCAAAATTTGCTGCTGACGGA 1259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1003 attaacgtgactgaagaagaaatataccaactgagactcgtgaactctgctgcgagcga 1122
QY 1260 TGGGGAACCAACAACATATAGTGGGTGCTTTAACGAGGTCAAAACMAAGTGTGCTCATAT 1319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1103 tgggaagacaaacactataaagtgtgtctttaaagaggttcaacaaatgtgtccatcat 1182
QY 1300 CCACACAGTCCGCTTTGGGGCCCTCTGCAGCTCAGACACTAGACAGAGCTGCCAAATGAC 1379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1183 ccaacagctgcttggggccctctgcagctcaagaactagagagcgtgtccaaatgac 1242
QY 1360 AGGAGCTTACAGCATATGCTGATGATCAAGTTCCAGAACATGGCCCATTTGATGCTTT 1439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1203 aggaagtttacagacaatagcttcaagatcaaatgataacaatgtgccatltgactt 1302
QY 1400 TGGGGCCCTTTTCAACAGAAATGAGAGTGTCTCTCAGCGCTCCATCCAGTTGAGATGA 1499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1303 tggggcccttccatcaggaatgagcgtctctcagcgtccatccagcttggagataa 1362
QY 1500 GGGATTTAACCTCCAGAACACGCACTGATGATGACACAGTATCGTGGACACGACCTG 1559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1303 tggggcccttccatcaggaatgagcgtctctcagcgtccatccagcttggagataa 1362
QY 1580 GGGAAAGACACTTTGTTCTTATACCTTGAGCAACGAGGCTCCCAATCTTCTGTG 1619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1403 gggaaagacacttgttcttctatacctggaacagcagcctcccaaatcccttctctg 1482
QY 1620 GGATCCAGTGGACAGAAAGCAAGGTGCTTTGTAGTGACAAAACACAAAATGCGCTTA 1679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1483 ggaatccagtgacagaagaagcaggtgttctgtgagcaaaaacacaaaatgacctta 1542
QY 1680 CCTCCAAATCCCAAGCATTTGCTTAAGTTGGCATTTGGAATTTACACTGTGCAAGCAAGCTC 1739
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1503 cctccaatcccaagcaltgtttaaagtlgacacttggaataacagctgtgcaagaagctt 1602
QY 1700 ACAAACTTTGACCCGCTGACGTGCTCCGCTGCGTCAATCTTACCCCTCCCAATTAC 1799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1603 acaaaccttgacactgactgtatcagatgccgtgtggttccaaatgtaaccttccaattac 1662
QY 1800 AGTGACTTCCAAACGAACAAGAGACACGCAAAATTCGCCAGCCCTGTGATTTATGC 1859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1603 agtgacttccaagaacgaagacacagcaaatctccagccctctgtgattatgc 1722
QY 1880 AAATTTGCCAAGAGGCTCCCAATCTTCAGAGGCGCACTGTACACGCCCTGATTTGATTC 1919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1703 aaatatctgcaagagcctcccaattctcagggccagtgltcaacagccctgattgaatc 1782
QY 1900 AGTGAATGGAAGAAACAGTTACTTGGAACTACTGATTAATGGAGCAGGTGCTGATGTAC 1979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1703 agtgaatggaagaaacagttactcttgaactactgtgataatgagcaggtgtctgtatgcac 1842
QY 1900 TAAAGATGAGGCTGTCTACTCAAGGTATTTCAAACTTATGACAGAAATGGTATGATCAG 2039
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1803 taagatgacaggtgttactactcaagttattcaacaactatgacaagaatgtgatatagcag 1902
QY 2000 TGTAAAGTGGGCGCTCTGGGAGAGGATTAACGCGACCGACAGGAGAGATGATACCCGACA 2099
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1903 tglataaagtggggtctcgggaagaggtlaacgcagccagagagagatglatccccagca 1962
QY 2100 GAGTGAAGACACTGTATACATACCTGCTGATGATGAGATGATGATGATGATGATGATGATG 2159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1903 gagtgaagcactgtatcaactactgtctgtgaattgagaatgataaataatgaaatccacc 2022
QY 2150 AAGACCTGAATTAATAAGATGATGTTCAACACAGCAAGTGTGTTTCACAGACAATC 2219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2003 aagacctgaataataaagtgatgttcaacaacaagaagtggtttcagcagaacatc 2082
QY 2200 CTCGGAGAGCTATTTGTGGCTTCTGATGTCGCAATGCTCCCATACTGATCTCTTCCC 2279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2083 ctgggaagctcatttctgtctctgactgttcccaaatgtctccataactgtactcttccc 2142
```


138 P-ESDB: AAB73715.
 139
 140 New antisense nucleotide, useful for treatment and prevention of
 141 bronchial asthma and chronic obstructive pulmonary disease
 142
 143 Example 1, Page 89-91; 104pp; Japanese.
 144
 145 The invention relates to an antisense nucleotide targeted to the mouse
 146 Gob-5 gene (coding sequence shown in AAB46101) or its human counterpart,
 147 the CLCA1 gene (coding sequence shown in AAB46102). The invention also
 148 relates to an antibody specific for the Gob-5 protein, medical and
 149 diagnostic compositions containing the antisense nucleotide or the
 150 antibody, and methods and kits for screening for compounds which inhibit
 151 the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 152 The antisense oligonucleotides and antibody are therefore useful for the
 153 treatment and prevention of bronchial asthma and chronic obstructive
 154 pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
 155
 156 Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;
 157
 158 Query: March 62.0%; Score 1743; DB 22; Length 2843;
 159 Best Local Similarity 77.8%; Pred. No. 0;
 160 Matches 2161; Conservative 0; Mismatches 595; Indels 21; Gaps
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609

742 GCACACATGTTGATCTTCTAGTGAATTCCTGACAGAAACCAACCAAGAAAGCT 801
 735 aacccaaatatactctctgcttgatctctgtacagaataatccatcaagaagcc 794
 802 CCAAAACAGCAAAATCAAAAATGCAATCTCCGACAGACACTGGGAAGTATCCGTATTC 861
 795 ccaaatgacaacaacaaagatgcaatctccgaagcagctggaagatccatcaagaatc 854
 863 GAGACCTTTAAGAAAACCACTCTATGACACAGACCCCAAAATCCCACTTTCATTC 921
 855 gagagcttcaagcaaacacccctccatgacagccacccgcgaacccactctccactg 914
 922 CTGCGATTTGGCAAGAATTTGTGTTTATGCTTGGACAAATCTGGAAAGATGGCGACT 981
 915 ctgcacaaattgacaaagaattgtgtcttaagctctgtaagctccggagcagctgaac 974
 982 GGTAAACCGCTCAATGCACTGCAATGCAAGCAGCGAGCTTTCTCTGCTGACAGCTGAG 1041
 975 gatgctcttcaaccgaatgaaatcagaagcagcgtcttccgctgcaagctgtgag 1034
 1042 CTGGGGCTCTGGGTTGGATGGTGAATTTGACAGTGTGCCCATGTACAAAGTGAATCT 1101
 1035 caggatctcctggtcggatggtgacgtctgacagctgctgctatgataaagcgaactc 1094
 1102 ATACGATTAACAGTGGCAGTACAGGAGACACTGCCAAAAGATTACCTGACAGACT 1161
 1095 aacacgttaaacagctgtgtcagacagatctgcatcaagcacttcccaagatct 1154
 1162 TCAGAGGAGGACGTCATGCACTGCAAGCGGCTTGCATGGCACTTACTGTATGGAAGAA 1221
 1155 gcagagaggaactctatctgctcgtcctcgacagacatctacagctgataaagaag 1214
 1222 TATCCAACTGATGATCTGAAATTTGCTGCTGACGATGGGAGAGAAACACTTAACT 1281
 1215 tatcaactgtagctcgaatgtgtcgtaccgagtggagagaaacacactgagc 1274
 1282 GGGTCTTTAAAGAGTCAAAACAAAGGTGCTGCATCATCACAAGTGTGTTGGGGCC 1341
 1275 agctcttctgacgtgtgaagcagagcgggacatccatcaagctgagccctgagacg 1334
 1342 TCTGAGCTCAAGACTAGAGAGCTGTCGCAAAATGACAGAGGTTTACAGCATATGCT 1401
 1335 gctgcgcttaaaagagctgagcagctgccaacagcagagggccctgacagcaactct 1394
 1402 TCAGATCAAGTTGACAAATGCTCATTTGATGCTTTGGGCGCCCTTTCATCAGGAAT 1461
 1395 tcgagctcagctcagaacaaatgctgtgtgcttccgacactctcccgagaaat 1454
 1462 GGAGCTGCTCAGCGCTTCATCCAGCTTGAAGTGAAGGATTTAACTCCAGAAAGC 1521
 1455 gcggagatcgctcagcactccatccagctgagcagggaggttaactccagataac 1514
 1522 CAGTGGATGAATGCGACAGTATCGTGGACAGCAGCTGGGAAAGACACTTGTGTTCT 1581
 1515 caatgtagatgctgctgagctgagctgagcagctcgtgagcagaaagaaactgttctt 1574
 1582 ATCAGCTGACACGACGCTCCCAAAATCCTCTCTGGATCCCAAGTGAAGCAAGCA 1641
 1575 atccctctgaaacgcacccctccaaatattatctgagctccagagaggtgaaacaa 1634
 1642 GGTGGCTTTGATGAGCAAAACACCAAAATGGCTACCTCCCAAAATCCCGAGCTTCT 1701
 1635 aatgatttcaatcaac 1694
 1702 AAGTGTGCACTGGAATTAAGTCTGCAAGCAAGCTCACAACCTTGAACCTGACTCTC 1761
 1695 aaggtgtcttcttgaaataacacacacacacacacacacacacacacacacacacac 1754
 1762 ACGTCCGCTGCTCAATGCTACCTGCTCAATTTACAGTGAAGTCAAGCAAGCAAG 1821
 1755 aactccctgacgaagcagctgacacacacacacacacacacacacacacacacacac 1814
 1822 GACACGACGAATTCGCCAGCCCTCTGATTTATGCAAAATTTCCGCAAGGAGCCCTCC 1881

1815 aacacaggaataatcccccagccctgtaacagctgtagcaagatctccgaagagccctg 1874
 1882 CCAATTTCCAGGCGCCAGTCTCAGACCCCTGATTTGAATCAGTAATGGAAGAAACATTACC 1941
 1875 cctattctcaagccagcctcagcctctgattgataatctgtaagtaagaaacagtaac 1934
 1942 TTGAACTACTGATTAATGAGCAGAGTGTGCTGATGCTTAAAGATGAGGTGTACACA 2001
 1935 ctggaattactcagtaaaagagcagctgctgacacacacacacacacacacacacacac 1994
 2002 AGTATTTCAAACTTATGACAGATGATGATGATGATGATGATGATGATGATGATGATG 2061
 1995 aggttttcaagcttctgtagtaagctgtagtaagctgtagtaagctgtagtaagctgtag 2054
 2063 GAGTTTAAACGCGCAGCAGGAGTGTATACCCAGCAGAGTGGAGCTGATGATGATGAT 2121
 2055 ggaatcaactcagacagagagcagacaccccgaaacacacacacacacacacacacacac 2114
 2132 GCTGGAATGAGAAATGATGAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 2181
 2115 ggtctgattgagatgagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 2169
 2132 GATGTTCAACACAGCAAGTGTGTTTACAGCAGACATCTCGGAGGCTCATTTGTGCT 2241
 2170 -atgtcaagaacagcagctgtctcagcagacatctcagagagctgtctgtgctg 2228
 2242 TCTGATGTC--CAATAGCTCCATACCTGATCTCTCCACCTGGCAATATCCAC 2298
 2229 accaatgtcccccagcagctccatctccatctccatctccatctccatctccatctccat 2288
 2239 CTGAAGGCGGAATTCACGCGGCGAGCTCTATTAATGATGATGATGATGATGATGATGAT 2358
 2249 ctgaagcagcagatccaaagcagacactggtgaatctgaagctggaagcgtccctggagat 2348
 2359 GATTATGACCATGGAACGCTCACAAGTATATCATTTGAATGAAGTACAAATTTCTTAT 2418
 2339 gactcagacacagagagcgtctccacacacacacacacacacacacacacacacacacac 2408
 2439 CTCAGACAGTTCATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2478
 2439 ctcaagagac 2468
 2439 GCAACCTGAGGAAGCTTTTGTGTTTAAACAGAAACATTTACTTTGAAATGAGCACA 2538
 2469 gccagctctgaggaatcttggatgtaactgtaggagcacaactcttggaatgagcaca 2528
 2539 GATCTTTTCAATGCTATTCAGGCTGTTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2598
 2539 gatctctcattgtcatccagcctgtgtgataagctcgaatcgaatcgaatcgaatcgaatc 2588
 2539 ATTGACAGAGTATCTTTTATTTATTCCTCCACAGACTCCGACAGACCAATGCTGAT 2658
 2589 atgcaagcagcgtctgctgtatccatcccgctcag-----gagccgacatcccgaa 2639
 2639 GAAACGTTGCTCTCTGCTTAATTTATTCATCAACAGACACCATCTCTGGCATTCACAT 2718
 2670 gactcaactcccccctctgctgacatcagcatcaacacacacacacacacacacacacac 2699
 2719 TTAATAATTAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2778
 2730 ctgaagaataatgagatgagctgaggaagaaatgcaagctgacacacacacacacacacac 2759
 2779 TTTGTCAGATGAATATAA 2795
 2790 ttcagcagaagaatcaaa 2776

RESULT: 11
 AAF81925
 ID AAF81925 standard; cDNA; 2931 BP.
 XX
 AC AAF81925;

PR	24-JUN-1998	98US-0090472
PR	24-JUN-1998	98US-0090535
PR	24-JUN-1998	98US-0090538
PR	24-JUN-1998	98US-0090540
PR	24-JUN-1998	98US-0090547
PR	25-JUN-1998	98US-0090678
PR	25-JUN-1998	98US-0090676
PR	25-JUN-1998	98US-0090688
PR	25-JUN-1998	98US-0090690
PR	25-JUN-1998	98US-0090694
PR	25-JUN-1998	98US-0090695
PR	25-JUN-1998	98US-0090696
PR	26-JUN-1998	98US-0090662
PR	26-JUN-1998	98US-0090663
PR	01-JUL-1998	98US-0091358
PR	01-JUL-1998	98US-0091360
PR	01-JUL-1998	98US-0091374
PR	02-JUL-1998	98US-0091444
PR	02-JUL-1998	98US-0091486
PR	02-JUL-1998	98US-0091519
PR	02-JUL-1998	98US-0091626
PR	02-JUL-1998	98US-0091628
PR	02-JUL-1998	98US-0091633
PR	02-JUL-1998	98US-0091646
PR	02-JUL-1998	98US-0091673
PR	07-JUL-1998	98US-0091978
PR	07-JUL-1998	98US-0091982
PR	09-JUL-1998	98US-0092182
PR	10-JUL-1998	98US-0092472
PR	10-JUL-1998	98US-0093435
PR	30-JUL-1998	98US-0094651
PR	04-AUG-1998	98US-0095282
PR	04-AUG-1998	98US-0095285
PR	04-AUG-1998	98US-0095302
PR	04-AUG-1998	98US-0095318
PR	04-AUG-1998	98US-0095321
PR	04-AUG-1998	98US-0095325
PR	10-AUG-1998	98US-0095519
PR	10-AUG-1998	98US-0095529
PR	10-AUG-1998	98US-0096012
PR	11-AUG-1998	98US-0096143
PR	11-AUG-1998	98US-0096146
PR	12-AUG-1998	98US-0096329
PR	17-AUG-1998	98US-0096757
PR	17-AUG-1998	98US-0096766
PR	17-AUG-1998	98US-0096773
PR	17-AUG-1998	98US-0096791
PR	17-AUG-1998	98US-0096667
PR	17-AUG-1998	98US-0096691
PR	17-AUG-1998	98US-0096694
PR	17-AUG-1998	98US-0096893
PR	17-AUG-1998	98US-0096897
PR	18-AUG-1998	98US-0096949
PR	18-AUG-1998	98US-0096950
PR	18-AUG-1998	98US-0096959
PR	18-AUG-1998	98US-0096660
PR	18-AUG-1998	98US-0097022
PR	19-AUG-1998	98US-0097141
PR	20-AUG-1998	98US-0097218
PR	24-AUG-1998	98US-0097661
PR	26-AUG-1998	98US-0097851
PR	26-AUG-1998	98US-0097852
PR	26-AUG-1998	98US-0097854
PR	26-AUG-1998	98US-0097955
PR	26-AUG-1998	98US-0097971
PR	26-AUG-1998	98US-0097874
PR	26-AUG-1998	98US-0097878
PR	26-AUG-1998	98US-0097879
PR	26-AUG-1998	98US-0097886
PR	26-AUG-1998	98US-0098014

31 AUG-1998; 98US-0098525.
16 SEP-1998; 98US-0100634.
12 JAN-1999; 99US-0115565.
RA (GEPH) GENENTECH INC.
Bayer K., Chen J., Goddard A., Gurney AL, Smith V, Watanabe CK;
Wong WL, Yuan J;
Weil; 2000-072883/06.
P.FSDB; AA566749.

Membrane-bound proteins and related nucleotide sequences

Claim 2; Fig 273; 822pp; English.

The invention provides membrane-bound PRO polypeptides and polynucleotides encoding them. The PRO sequences of the invention were identified based on extracellular domain homology screening. The PRO sequences have homology with proteins including LDL receptors, TIE ligands and various enzymes. The membrane-bound proteins and receptor molecules are useful as pharmaceutical and diagnostic agents. Receptor immunodessins, for instance, can be used as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. The PRO encoding sequences are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques.

Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;

Query Match 46.4%; Score 1304; DB 21; Length 3265;
Best Local Similarity 69.8%; Freq. No. 0;
Matches 1869; Conservative 0; Mismatches 785; Indels 24; Gaps

1 ACCAATGGGGCCATTAAAGAGTTCTGTGTCCTTGAATTCACCTTCTGAAGGGGC 80
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2 AACCAATGGGATCACCAGAAGTCTTGTTCTCTCTCTAGTCTGCTCGTGGACCAAG 80
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3 CCTGATTAATTCATCATTCACCTGACCAACAACATGGCTATGAGGACATTCGTTCAAT 140
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4 ---aaattcttcctcatlaagcagaatgaataaaggcttggaagtattgcatctgttat 137
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5 CGACCCCAATGTGCCAGAAAGATGAACACTCTATTACAAATAAAGACATGTGATACCCA 200
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6 agatccatagttgcacagaagaatgaaaaaaataatgaacaatatagaglatatggtactac 197
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7 GGCAATCTCGTATCTGTTGTAAGCTAACAGGAAGGATTTTTCAAAAGATGTTGCAT 260
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
8 acgtcttaccgtacctgttagaacgccacagaataaatctllttlccaataaagatcatat 257
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9 TTGTGATTCCTGTAACATGACACACAAAGGCTGACTATGTGAGACCAAAATCGAGACTA 320
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
10 ataattctctggaattggaagaaaatctcaagtaacaaaaggccaataaatgaanaacca 317
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
11 CAATAATGCTGATGTTCCTGTTGCTGAGTCTACTCTCCAGSTAAATGATGAACCTTACAC 380
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
12 taacaatcgatgattatagttcaccaactacatccocaggtagatgaacacatacac 377
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13 TGAGAGATGGGCACTGTGGAGACAAAGGATGAAGATCCACCTCACCTCCGATTTTCAT 440
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
14 caagcagttccacagaaagtggagagaagaagcgaaatacatactcaatccaccctgacctct 437
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15 TCACAGAAAAAAGTTAGCTGAATATGACACCAACAGATGGGATGTCATAGATGGGAC 500
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
16 acttggaataaaaaaataatgatatggaacaccacaggcaaatcgtttgccaatgagatgggc 497
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 TCATCTAGCATGGGAGATTTGACGAGTACATATATGATGAGAAATTCATCTACTTACAA 560
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
18 tcaactccggttggaagtgtttgaatgaatgaacataaataacagctctcttaccgtgctaca 557
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
AA92092
ID AAF92092 standard; cDNA; 3265 BP.
AC AAF92092;
DT 15-MAY-2001 (first entry)
XX
DE Human PRO1124 cDNA.
XX
KW Human; PRO protein; mapping; ss.
OS Homo sapiens.
XX
PN WO200116318-A2.
XX
PD 08-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-US23328.
XX
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 07-DEC-1999; 99US-0169495.
PR 09-DEC-1999; 99US-0170262.
PR 11-JAN-2000; 2000US-0175481.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 01-MAR-2000; 2000WO-US05601.
PR 03-MAR-2000; 2000US-0187202.
PR 25-APR-2000; 2000US-0193397.
PR 22-MAY-2000; 2000WO-US14042.
PR 05-JUN-2000; 2000US-0209832.
XX
PA (GENE) GENENTECH INC.
XX
PI Eaton DL, Filvaroff E, Gerlitsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2001-183260/18.
XX
DR P-PSDB; AAB87560.
XX
PT Eighty four nucleic acids encoding PRO polypeptides, useful in
PT molecular biology, including use as hybridization probes, and in
PT chromosome and gene mapping.
XX
PS Claim 2; Fig 69; 27bpb; English.
XX
CC The present sequence is the coding sequence for a human PRO polypeptide
CC (secreted and transmembrane). The PRO protein, and PRO agonists, PRO
CC antagonists or anti-PRO antibodies are useful for preparation of a
CC medicament useful in the treatment of a condition which is responsive to
CC the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
CC protein may also be employed as molecular weight markers for protein
CC electrophoresis. The PRO coding sequence has applications in molecular
CC biology, including use as hybridisation probes, and in chromosome and
CC gene mapping.
XX
SQ Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;

Query Match 46.4%; Score 1304; DB 22; Length 3265;
Best Local Similarity 69.8%; Pred. No. 0;
Matches 1869; Conservative 0; Mismatches 785; Indels 24; Gaps 7;

QY 21 ACCATGGGGCCATTAAAGATTCTGTCATTGATTTCTACCTCTAGAAGGGC 80
DB 21 acaaatggggccattaaagattctgttcctccttaagctcgtgcgcacacagtc 80
QY 81 CCTGATTAATTCATTCATTCAGTGAACAATGCTGATGAAGCATTTGCTTGCAT 140
DB 81 ---aaatctctcctcatcagctgaataaataatgagcttggaagatattgtctgtat 137

CGACCCCAATGTCCAGAAAGATGAAGAACTCATTTCAACAAATTAAGACATGGTACCCA 200
118 agatcctgagtgccagaagatgtaaaaaataatgtaacaataagagatattgtgctac 197
201 GGCATCTCTGATCTGTTGAGCTGACAGAAAGCATTTTATTTCAAAATGTTCCCAT 260
158 agcttcgagctcgttctggaagccacagaaaaaagatlltlltcaaaaatglatcctat 257
261 TTTGATTTCTGAAACATGGAAGACAAAGCTGACATTTGTGACACAAATTTGAGACTA 320
158 attaatctctggaatattggaagaaatactcagtaacaaagccaaacaaagaaacaa 317
321 CAAAAATGCTGATGTTCTGCTGCTGATCTACTCTCCAGGATATGATGAACCTTACAC 380
318 taacatgctgattatagttcacacacttaacccacaggaagagatgaacataacac 377
381 TGACGAGATGGGCACTGTGGAGAGAGAGGTGAAGATCCACCTCCATCTGATTTCAAT 440
388 caagcagcttcaacagaaatggaagaaagcgaataacatcactcaccctcgaacctct 437
411 TGCAGGAAAAAGTTAGCTGAATATGACACACAGCTAGGCGCATTTGTCCATGATGGGC 500
418 acttgaaataaaataaataatgataatgacacacagcaacatgcttgcacatgagtgagg 497
501 TCATCTAGCATGGGAGATATTGACGAGTACAAATAATGATGAAATTTCTACTTATCAAA 560
498 tcaccccggtggygagtggttattgatgaataagagctcacccttaccgctgctaa 557
561 ---TGGAAGATATGACAGCTAAGATGTTGACAGAGTATTACTGTGACAAATTAATAA 617
568 gtcaaaaaaaatcgaagcaacaaaggtctccgaaaglatcctcgtgagaataagagltta 617
618 GAAGTTCAGGAGGAGCGAGCTGTTTACACAAAAGATCCACATTCCTAAATTAAGTAA 677
618 taagtgcagaagagagagctgtctttagtagatgagatgacgaatgattcacaacaaact 677
678 CTATGAAAAGATGATGAGTTGTTGTTCTCAATCCCGCAGAGGAGGAGGCTTTATTAAT 737
678 gtatggaagaagattgtcaattcttcctcgtataagaacaaacagaataacacataat 737
738 GTTTCGCAACAACTGTTGATTTCTATGTAATCTGTACAGAAACAAACCAACAAAGA 797
738 gtltatgcaaaagatgattcgttctgtaattgtaacgaaaaaacacataaataa 797
738 AGCTCCAAACAAAGCAAAATCAAAATCAATCTCCAGACCAATGGAAGTGCCTGA 857
738 agctccaagcctcacaataacataaagtgcaatttgaagtaactggygagtgatgca 857
858 TTCTGAGACTTTAAGAAAACCACTCTTATGACACACAGCCCAAAATCCACTTCTC 917
858 tttctgaggttttaaaaaacacataccatggtgacacacccctccacactgtctctc 917
918 ATTGCTGAGATGAGCAAAAGAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 977
918 atgtcgaagatcagtaacaaagattggtttagttagttagttagttagttagttagttag 977
978 GACTGCTAACCGCTCAATTCAGATGATCAAGCAGCGCCAGCTTTTCTGTCGACAGAT 1037
978 ggttaagagccgccttaataatcgaatgatacaagaagaacaaacattctcgtgagagctg 1037
1038 TGAGCTGGGGCTCTGGGTTGGATGGTGCATTTGACATGCTGCTCCCATGTACAAAGTGA 1097
1038 tgaataatgacatcgtggtggtggtggttcaactttagttagttagttagttagttagt 1097
1098 ACATATACAGATTAACAGAGGAGTACAGAGGAGGACACATCCGCAAAAGATTACTCCAGC 1157
1098 gctaatcacaataaagaagatgatagaagaacacacacacacacacacacacacacacac 1157
1158 AGCTTACAGAGGAGGAGCTTCATTCAGCGGCTTTGATGCGCATTTACTGTATGAGAA 1217
1158 tctctcgtgagagagatcctcgtcgtgaaataatgacatcctcaggtggtgagga 1217
1218 GAAATAT---CCAACTGATGATCTGAATTTGCTGCTGACGAGATGGGGAAGACAAACAC 1274

Db 1578 cttctcacaacagcagctcgtccagatattctcttggatccagtggaac 1637
OY 1635 GAACGACGAGCTGCTTTGCTAGTGCACAAAACACCAAAATGGCTACCTCCAAATCCAGG 1694
Db 1638 aatcatggaanaattccacagtgatgcaacttccaaatgctcattcagatccagc 1697
OY 1695 CATTGCTAAGGTTGGCACTTGGAAATACAGTCTGC-----AGCAAGCTCACAAACTT 1748
Db 1698 aactgcaagcttggcacttggcattacaacttccaagccaagcgaaacccaact 1757
OY 1749 GACCTGACGCTCAGGCTCCGCTGCTCAAGTACCTGCTCCAAATACAGTGAATTC 1808
Db 1758 aactatcacagtaacttcgagcagaatattctctgctccatcacagtgatgc 1817
OY 1809 CAAGGACGACGACACACCAAAATCCCGAGCTCTGCTAGTATGCAAAATATTCG 1868
Db 1818 taatatgaataagcagtaaacagcttcccaagccaatgatgtttacgcagaatctc 1877
OY 1869 CCAAGGAGCTCCCAATTCACAGGCGCAGTGCACAGCCCTGATTGATCATGTAATGG 1928
Db 1878 acaaggatattacccttctcgtcgaacagtgatgcttcatctgaatcacagaatgg 1937
OY 1929 AAAACAGTACTCTGTAAGTACTGATATGAGCAGGTCGTGATGCTACTAAGATGA 1988
Db 1938 acatacagaagcttggacacttggataatgctcagcgctgattcttccaagaatga 1997
OY 1989 CGGTGCTCTCAAGGATTTTACACTATGACACGAAAGTGTAGATACAGTGAAGT 2048
Db 1998 tggagctcactccaggtatttacaagcatatcagaanaatggcagatataagcttaaaagt 2057
OY 2049 GCGGCTCTGGGAGGATTAAAGCAGCAGCAGGAGAGTGTATACCCAGCAGAGTGGAGC 2108
Db 2058 tcgggctatggagagcagaacacagtcagcgtataatcagcctccatcgaatagagc 2117
OY 2109 ACTGTACATPACTGCTGCTGATGAGATGATGAATACAAATGGAATCCACCAAGACTGA 2168
Db 2118 cgcgtacataccagcctgggtagtgacgaggaatgaacgaacccgcgaagacclga 2177
OY 2169 AATTAAATAGGATGATGTTCAACACAGCAAGTGTGTTTACAGCAATCATCTCGGAGG 2228
Db 2178 aattgat--gagatactcagacacacacttggagatcttcagccgaacagcatccggag 2234
OY 2229 CTATTTTGGCTTGTGATGTCCAAAATGCTCCATACCTGATCTTTCACACTGGCCA 2288
Db 2235 tgcatttggatatacacaagctcccaagccttcctcgtgacataccaccaagtlca 2294
OY 2288 AATCAACGACCTGAAGCCGGAATTCACGCGGCGAGTCTCATATTCGACTTGGACAGC 2348
Db 2295 aatcacagaacttgaatgacacagtlcagag--ataagattatcttacaatgagcagc 2351
OY 2349 TCCCTGGGATGATTTGATGACATGAGCAAGCTCACAAGTATATTCGATTAAGTACAG 2408
Db 2352 accagagagtaatttggatgaggaagttcaacgttatatcatagaataagtgcaag 2411
OY 2409 TATTTCTGATCTCAGACAGCACTTAATGATCTCTCAAGTGAATFACACTGCTCAT 2468
Db 2412 tatctcgtatccaagagacagtttggatgctctccaagtaataactactgactctgc 2471
OY 2469 CCCAAAGGAGCAAGCTGAGCAAGTCTTTTGTAAACAGCAAAACATTTACTTTTGA 2528
Db 2472 accaaagagggccaactccaagaaagccttgcatttaacccaagaataatctcagaaga 2531
OY 2529 AATGGCACAGATCTTTTTCATTTGCTATTGACGCTGTTGATTAAGTCGATCTGAATCAGA 2588
Db 2532 aatgcaaccacataatttattgcatataaagtagataaagaatltgacatcaaa 2591
OY 2589 AATATCCACATTTGCAAGGATATCTTTGTTTATTCCTCCACAGACTCCGCCAGAGA--C 2645
Db 2592 agtatccaacattgcaacagtaacttggatccctcaagcaaaacccgatgacatga 2651
OY 2646 ACCTAGTCCGATGAAGAGTGTGCTCTTGTCTTAATA 2683
Db 2652 tccataactactcctactcctactcctactcctgata 2689

Search completed: April 4, 2002, 03:21:09
Job time: 60421 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:08:35 ; Search time 22700.8 Seconds
(without alignments)
2044.269 Million cell updates/sec

Title: US-09-049-696-18
Perfect score: 2813
Sequence: 1 GAATATCAGCAGAGATGTAC.....AAATTAATCATTCATCCTTA 2813

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sy:*
27: em_un:*
28: em_vl:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htgo_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2807.2	99.8	3311	6	AX193489	AX193489 Sequence
2	2807.2	99.8	3311	9	AF039400	AF039400 Homo sapi
3	2790.4	99.2	2826	9	AF127036	AF127036 Homo sapi
4	2000.4	71.1	2022	9	AK024970	AK024970 Homo sapi
5	1898.4	67.5	3079	4	AF095584	AF095584 Sus scrofa
6	1743.2	62.0	2937	10	AB017156	AB017156 Mus muscu
7	1310.2	46.6	3204	9	AF127035	AF127035 Homo sapi
8	1304	46.4	3221	9	AK000072	AK000072 Homo sapi
9	1304	46.4	3265	6	AX092338	AX092338 Sequence
10	933.4	33.2	2984	4	BT036445	BT036445 Bos tauru
11	900.2	32.0	3317	4	AF001261	AF001261 Bos tauru
12	865.4	30.8	2765	10	AF115852	AF115852 Mus muscu
13	865.4	30.8	3137	10	BC008147	BC008147 Mus muscu
14	863.8	30.7	3058	10	AF108501	AF108501 Mus muscu
15	863	30.7	3022	10	AF047838	AF047838 Mus muscu
16	831	30.7	3471	10	AF052746	AF052746 Mus muscu
17	831	29.5	3415	9	AF043976	AF043976 Homo sapi
18	806	28.7	2820	4	AF001263	AF001263 Bos tauru
19	800.6	28.5	3288	4	AF001262	AF001262 Bos tauru
20	780.8	27.8	878	6	I95746	I95746 Sequence 8
21	672.2	23.9	1895	9	AK000138	AK000138 Homo sapi
22	553.8	19.7	2832	9	AF127980	AF127980 Homo sapi
23	553.8	19.7	4077	6	AX054697	AX054697 Sequence
24	552.2	19.6	2970	9	AF043977	AF043977 Homo sapi
25	550.6	19.6	3604	9	AB026833	AB026833 Homo sapi
26	437	15.5	113764	9	HS7651E10	HS7651E10
27	437	15.5	164891	2	AL358950	AL358950 Homo sapi
28	435.4	15.5	35278	9	AF039401	AF039401 Homo sapi
29	381	13.5	1265	4	AF001264	AF001264 Bos tauru
30	323.8	11.5	401	6	AX192467	AX192467 Sequence
31	187	6.6	124096	2	AL354988	AL354988 Homo sapi
32	163	5.8	140718	2	AL356270	AL356270 Homo sapi
33	151.6	5.4	375	6	AX071171	AX071171 Sequence
34	131.8	4.7	140718	2	AL356270	AL356270 Homo sapi
35	108.4	3.9	175591	2	AC068071	AC068071 Homo sapi
36	91.2	3.2	350	6	AX188011	AX188011 Sequence
37	91.2	3.2	454	6	AX185116	AX185116 Sequence
38	88	3.1	549	6	AX054715	AX054715 Sequence
39	73	2.6	307	6	AX187657	AX187657 Sequence
40	73	2.6	333	6	AX186584	AX186584 Sequence
41	73	2.6	423	6	AX184779	AX184779 Sequence
42	68	2.4	227	6	AX186451	AX186451 Sequence
43	52	1.8	7218	6	I66494	I66494 Sequence 14
44	47.8	1.7	110	10	AF077303	AF077303 Rattus no
45	46.4	1.6	392	11	G61053	G61053 SHGC-85269

ALIGNMENTS

RESULT 1
AX193489 3311 bp DNA PAT 15-AUG-2001
LOCUS: Sequence 1056 from Patent W00149716.
DEFINITION: AX193489
ACCESSION: AX193489
VERSION: AX193489.1 GI:15211440

KEYWORDS

SOURCE: human;
ORGANISM: Homo sapiens

REFERENCE: Xu, J., Lodes, M. J., Secrist, H., Benson, D. R., Meagher, M. J.,
Stolk, J. A., King, G. E., Wang, T., and Jiang, Y.
Compounds for immunotherapy and diagnosis of colon cancer and
Patent: WO 0149716-A 1056 12-JUL-2001.

AUTHORS

TITLE: JOURNAL
CORIXA CORPORATION (US)

FEATURES: Location/Qualifiers
source I. 3311

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 99.8%; Score 2807.2; DB 6; Length 3311;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2809; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GAAATCACAGGAGATGATACAGATGGGCGCATTTAGAGCTGTGTCTCTTGTAT 60
DB 328 GGAATCACAGGAGATGATACAGATGGGCGCATTTAGAGCTGTGTCTCTTGTAT 387
OY 61 CTTCACCTTCTAGAGGGGCGCTGAGTAATTCATCTACCTGACCTGACCAATGGCTAT 120
DB 388 CTTCACCTTCTAGAGGGGCGCTGAGTAATTCATCTACCTGACCTGACCAATGGCTAT 447
OY 121 GAAGCATTTGCTGTCGATGCAATGACCCCAATGTGCCAGAAATGAAACACTATTCACAA 180
DB 448 GAAGCATTTGCTGTCGATGCAATGACCCCAATGTGCCAGAAATGAAACACTATTCACAA 507
OY 181 AATAAGCATGTCGTCACGACGATCTGTATCTGTGTGAAAGCTACAGAAAGGATTT 240
DB 508 AATAAGCATGTCGTCACGACGATCTGTATCTGTGTGAAAGCTACAGAAAGGATTT 567
OY 241 TATTTCAAAAATGTTGCCATTTTGAATCTCTGAAACATGGAAGACAAAGGCTGATATGTG 300
DB 568 TATTTCAAAAATGTTGCCATTTTGAATCTCTGAAACATGGAAGACAAAGGCTGATATGTG 627
OY 301 AGACCAAACTTGAGACCTACAAAATGCTGATGTTCTGCTGCTGATCTACTCTCA 360
DB 628 AGACCAAACTTGAGACCTACAAAATGCTGATGTTCTGCTGCTGATCTACTCTCA 687
OY 361 GGTATGATGAAACCTTACACTGAGAGATGGGCACTGGGAGAGAAAGGAAAGATC 420
DB 688 GGTATGATGAAACCTTACACTGAGAGATGGGCACTGGGAGAGAAAGGAAAGATC 747
OY 421 CACCTGACTGCTGATTTGATGCAAGAAAAAGTTAGCTGAATATGACCAAGAGTAG 480
DB 748 CACCTGACTGCTGATTTGATGCAAGAAAAAGTTAGCTGAATATGACCAAGAGTAG 807
OY 481 GCATTTGTCAGATGAGTGGGCTCATCTAGATGGGAGATTTGACAGATACAAATATAT 540
DB 808 GCATTTGTCAGATGAGTGGGCTCATCTAGATGGGAGATTTGACAGATACAAATATAT 867
OY 541 GAGAAATCTCTATCTCAATGAGAAATACAGAGTAAAGTGTTCAGCAGGTATCT 600
DB 868 GAGAAATCTCTATCTCAATGAGAAATACAGAGTAAAGTGTTCAGCAGGTATCT 927
OY 601 GGTACAAATGTATGAAAGAGTGTGAGGAGGACGCTGTACACCAAAAAGATGACATTC 660
DB 928 GGTACAAATGTATGAAAGAGTGTGAGGAGGACGCTGTACACCAAAAAGATGACATTC 987
OY 661 AATAAGTAAGAGAGCTGTATGAAAAAGAGTGTGTTGTTCTCAATCCCGCAGAGC 720
DB 988 AATAAGTAAGAGAGCTGTATGAAAAAGAGTGTGTTGTTCTCAATCCCGCAGAGC 1047
OY 721 GAGAAAGCTTCTATATGTTGTCACACATGTTGATTTCTATAGTTGAATTCGTACAAA 780
DB 1048 GAGAAAGCTTCTATATGTTGTCACACATGTTGATTTCTATAGTTGAATTCGTACAAA 1107
OY 781 CAAAACCAACAAAGAGCTCCAAACCAAGCAAAATCAAAAATGGAATCTCCGAAGACA 840
DB 1108 CAAAACCAACAAAGAGCTCCAAACCAAGCAAAATCAAAAATGGAATCTCCGAAGACA 1167
OY 841 TGGGAAGATGCTGATGATTCGAGAGCTTTAAGAAACACTCTCTATGACAAACAGCA 900
DB 1168 TGGGAAGATGCTGATGATTCGAGAGCTTTAAGAAACACTCTCTATGACAAACAGCA 1227
OY 901 CCAATATCCACCTTCTCATGCTGAGATGACAAAGAAATTTGTGTTGATGCTCTTAC 960
DB 1228 CCAATATCCACCTTCTCATGCTGAGATGACAAAGAAATTTGTGTTGATGCTCTTAC 1287

OY 961 AAATCTGAGAGATGGGCACTGTGTACCGCCGCAATGCACTGAATCAACAGCCAGCTT 1020
DB 1288 AAATCTGAGAGATGGGCACTGTGTACCGCCGCAATGCACTGAATCAACAGCCAGCTT 1347
OY 1021 TTCTCTGTCAGACAGTGTAGCTGGGCTCAGGCTGTGGATGTGACATTTGACAGTCT 1080
DB 1348 TTCTCTGTCAGACAGTGTAGCTGGGCTCAGGCTGTGGATGTGACATTTGACAGTCT 1407
OY 1081 GCCCATGTACAAAGTGAAGTCTATACAGATTAACAGTGGCAGTGGACAGACACTGCC 1140
DB 1408 GCCCATGTACAAAGTGAAGTCTATACAGATTAACAGTGGCAGTGGACAGACACTGCC 1467
OY 1141 AAAAGATTAACCTGACGAGCTTCAGAGAGAGCTGCAATCTGAGGCGGCTGATGGCA 1200
DB 1468 AAAAGATTAACCTGACGAGCTTCAGAGAGAGCTGCAATCTGAGGCGGCTGATGGCA 1527
OY 1201 TTTACTGTGATTTAGAGAAATATCCAACTGATGATGTGAATTTGCTCTGACGGAT 1260
DB 1528 TTTACTGTGATTTAGAGAAATATCCAACTGATGATGTGAATTTGCTCTGACGGAT 1587
OY 1261 GGGGAAGCAACACTATTAAGTGGTCTTTAAGAGTCAAAACAAAGTGTGCATATC 1320
DB 1588 GGGGAAGCAACACTATTAAGTGGTCTTTAAGAGTCAAAACAAAGTGTGCATATC 1647
OY 1321 CACACAGTCCGCTTTGGGCGCTCTGACCTCAAGAACTAGAGAGCTGTCACAAATGACA 1380
DB 1648 CACACAGTCCGCTTTGGGCGCTCTGACCTCAAGAACTAGAGAGCTGTCACAAATGACA 1707
OY 1381 GGAGGTTTACAGACATATGCTTCAGATCAAGTTCAGAAATGCGCTCATATTGATCTTTT 1440
DB 1708 GGAGGTTTACAGACATATGCTTCAGATCAAGTTCAGAAATGCGCTCATATTGATCTTTT 1767
OY 1441 GGGGCGCTTTATACAGAAATGAGAGCTCTCTCTAGGCTCTCATTCACCTTGAGATAG 1500
DB 1768 GGGGCGCTTTATACAGAAATGAGAGCTCTCTCTAGGCTCTCATTCACCTTGAGATAG 1827
OY 1501 GGATTTAACCTTCAGAACAGCAGTGTGATGATGCAACAGTATGCTGAGACAGACGCTG 1560
DB 1828 GGATTTAACCTTCAGAACAGCAGTGTGATGATGCAACAGTATGCTGAGACAGACGCTG 1887
OY 1561 GGAAGAGCACTTTGTTCTTATACCTGAGCAACAGCAGCTCCCAATCTCTCTGCG 1620
DB 1888 GGAAGAGCACTTTGTTCTTATACCTGAGCAACAGCAGCTCCCAATCTCTCTGCG 1947
OY 1621 GATCCAGTGTGAGCAAGCAAGTGGGCTTTGATGGGCAAAAACCAAAATGGCCCTAC 1680
DB 1948 GATCCAGTGTGAGCAAGCAAGTGGGCTTTGATGGGCAAAAACCAAAATGGCCCTAC 2007
OY 1681 CTCCAAATCCAGAGCATTTGCTAAGTGTGCACTTGGAAATACAGTCTGCAAGAGCTCA 1740
DB 2008 CTCCAAATCCAGAGCATTTGCTAAGTGTGCACTTGGAAATACAGTCTGCAAGAGCTCA 2067
OY 1741 CAAACCTTGACCTGACTGTACGTCCTGCTGCTCAATGCTTACCTGCTCCAAATTACA 1800
DB 2068 CAAACCTTGACCTGACTGTACGTCCTGCTGCTCAATGCTTACCTGCTCCAAATTACA 2127
OY 1801 GTGACTTCCAAAGCAAGCAAGCAGCAAAATTTCCCAAGCCCTGCTGCTTATGCA 1860
DB 2128 GTGACTTCCAAAGCAAGCAAGCAGCAAAATTTCCCAAGCCCTGCTGCTTATGCA 2187
OY 1861 AATATTCGCCAAGAGAGCTCCCAATTTCTAGGSCCAAGTCAACAGCCCTGATTTGAATCA 1920
DB 2188 AATATTCGCCAAGAGAGCTCCCAATTTCTAGGSCCAAGTCAACAGCCCTGATTTGAATCA 2247
OY 1921 GTGAATGCAAAACAGTACTTGTGAAGTCTGATTAATGAGAGAGTGTGATGCTACT 1980
DB 2248 GTGAATGCAAAACAGTACTTGTGAAGTCTGATTAATGAGAGAGTGTGATGCTACT 2307
OY 1981 AAGGATGACGCTGTCTCTCAAGGATTTTCAACATATGACAGAAATGATGATATAGT 2040
DB 2308 AAGGATGACGCTGTCTCTCAAGGATTTTCAACATATGACAGAAATGATGATATAGT 2367

421 CACCTCACTCTGATTTTCATTCAGGAAAAAGTTAGTGAATATGACACAGTAGG 480
748 CACCTCACTCTGATTTTCATTCAGGAAAAAGTTAGTGAATATGACACAGTAGG 807
481 GCATTTGTCTGATGAGTGGGCTCATCTCAGATGGGAGATTTTGGACAGTACATAATGAT 540
808 GCATTTGTCTGATGAGTGGGCTCATCTCAGATGGGAGATTTTGGACAGTACATAATGAT 867
541 GAGAAATTTCTATTTCCATGAGAAATATCAAGCAGTGAAGTTCACAGGATTTACT 600
868 GAGAAATTTCTATTTCCATGAGAAATATCAAGCAGTGAAGTTCACAGGATTTACT 927
601 GGTACAAATGTAGTAAAGAGTGTCAAGGAGGAGCTGTTTACCAAAAAGATGACATTC 660
928 GGTACAAATGTAGTAAAGAGTGTCAAGGAGGAGCTGTTTACCAAAAAGATGACATTC 987
661 AATTAAGTACAGGAGTCTATATGAAAAAGATGTGAGTTTGTCTCCATCCCGCAGAG 720
988 AATTAAGTACAGGAGTCTATATGAAAAAGATGTGAGTTTGTCTCCATCCCGCAGAG 1047
721 GAGAGGCTTCTATATGTTTGCACACATGTTGATTTCTATGTTGAATTCGTACAGA 780
1048 GAGAGGCTTCTATATGTTTGCACACATGTTGATTTCTATGTTGAATTCGTACAGA 1107
781 CAAAACCAACAAAGAAAGCTCCAAACAAAGCAAAATTCAAAATTCGATCTCGAAGACA 840
1108 CAAAACCAACAAAGAAAGCTCCAAACAAAGCAAAATTCAAAATTCGATCTCGAAGACA 1167
841 TGGGAAGATCCGATTTCTGAGAGCTTTAAGAAACCACTCCATATGACAAACAGACA 900
1168 TGGGAAGATCCGATTTCTGAGAGCTTTAAGAAACCACTCCATATGACAAACAGACA 1227
901 CCAAAATCCACCTTCTCATTTGCTGAGATTTGACAAAGATTTGTTAGTCTCTGAC 960
1228 CCAAAATCCACCTTCTCATTTGCTGAGATTTGACAAAGATTTGTTAGTCTCTGAC 1287
961 AATTTGGAAGATGGGAGCTGTGTACCCGCTCAATTCGATGATCAAGAGGCGAGCTT 1020
1288 AATTTGGAAGATGGGAGCTGTGTACCCGCTCAATTCGATGATCAAGAGGCGAGCTT 1347
1021 TTCTGCTGAGAGAGTGTGAGCTGGGGCTCTGGGTTGGGATGGGATGAGACATGCT 1080
1348 TTCTGCTGAGAGAGTGTGAGCTGGGGCTCTGGGTTGGGATGGGATGAGACATGCT 1407
1081 GCCCATGTACAAAGTGAACCTCATACAGATAAACAGTGGCAGTGAAGGAGACACTCGCC 1140
1408 GCCCATGTACAAAGTGAACCTCATACAGATAAACAGTGGCAGTGAAGGAGACACTCGCC 1467
1141 AAAAGATTAAGTGAAGAGCTTTCAGAGGAGGAGCTTCATCTGACGAGGCTTCGATCGCA 1200
1468 AAAAGATTAAGTGAAGAGCTTTCAGAGGAGGAGCTTCATCTGACGAGGCTTCGATCGCA 1527
1201 TTACTGTGATTAGGAAGAAATATCAACTAGTGAATGTGAATTTGTCTGCTGAGGAT 1260
1528 TTACTGTGATTAGGAAGAAATATCAACTAGTGAATGTGAATTTGTCTGCTGAGGAT 1587
1261 GGGGAAGACAACTATATGAGTGGGCTGTTTAAAGAGTTCGAAAGAGTGGTGCATATC 1320
1588 GGGGAAGACAACTATATGAGTGGGCTGTTTAAAGAGTTCGAAAGAGTGGTGCATATC 1647
1321 CACACAGTGGTGGGGGCTCTGACGCTCAAGACATGAGAGGAGCTGTCCAAATGACA 1380
1648 CACACAGTGGTGGGGGCTCTGACGCTCAAGACATGAGAGGAGCTGTCCAAATGACA 1707
1381 GAGAGTTACAGACATATGCTTCAGATCAAGTTCAGAAATGAGGCTCATGATGCTTTT 1440
1708 GAGAGTTACAGACATATGCTTCAGATCAAGTTCAGAAATGAGGCTCATGATGCTTTT 1767
1441 GGGGGCTTTTCATGAGAAATGAGGCTGTCTCTAGGCTTCATTCAGCTTGAGAGTAAG 1500
1768 GGGGGCTTTTCATGAGAAATGAGGCTGTCTCTAGGCTTCATTCAGCTTGAGAGTAAG 1827
1501 GGATTAACCTTCAGAGAGCGAGTGAATGAGTGAAGTGTGAGACAGCCTG 1560
1838 GGATTAACCTTCAGAGAGCGAGTGAATGAGTGAAGTGTGAGACAGCCTG 1887
1561 GGAAGAGACATTTGTTCTTATACCTGAGACAGCAGGCTCCCAATCTCTCTGG 1620
1888 GGAAGAGACATTTGTTCTTATACCTGAGACAGCAGGCTCCCAATCTCTCTGG 1947
1621 GATCCAGTGGACAGACAGAGTGGCTTTGTAGTGAACAAAACCAAAATGGCTTAC 1680
1948 GATCCAGTGGACAGACAGAGTGGCTTTGTAGTGAACAAAACCAAAATGGCTTAC 2007
1681 CTCCAATCCAGAGGATGTGATGAGTGGCACTTGAATACAGTCTGACAGCAAGCTCA 1740
2008 CTCCAATCCAGAGGATGTGATGAGTGGCACTTGAATACAGTCTGACAGCAAGCTCA 2067
1741 CAACCTTGACCTGACGTGACAGTCCGCTGCTCCATGCTTACCTGCTCCCAATTTCA 1800
2068 CAACCTTGACCTGACGTGACAGTCCGCTGCTCCATGCTTACCTGCTCCCAATTTCA 2127
1801 GTGACTTCCAAAGCAAGACAGACAGCAATTTCCCAAGCTCTGTGTATGATCA 1860
2128 GTGACTTCCAAAGCAAGACAGACAGCAATTTCCCAAGCTCTGTGTATGATCA 2187
1861 AATATTGCCAAGAGAGCTCCCAATTTCTAGGGCAGTGTACAGGCTGATTTATGCA 1920
2188 AATATTGCCAAGAGAGCTCCCAATTTCTAGGGCAGTGTACAGGCTGATTTATGCA 2247
1921 GTGAATGGAAGAAACAGTTACCTTGGAACTGATGATTAATGAGAGAGTGTGATGCTACT 1980
2248 GTGAATGGAAGAAACAGTTACCTTGGAACTGATGATTAATGAGAGAGTGTGATGCTACT 2307
1981 AAGGATGACGGTGTCTACTCAAGGATTTTCAACTATATGACAGAAATGATATAGT 2040
2308 AAGGATGACGGTGTCTACTCAAGGATTTTCAACTATATGACAGAAATGATATAGT 2367
2041 GTAAAAGTGCAGGCTCTGGAGAGATTAAAGCAGCAGAGAGAGAGATATCCACAG 2100
2368 GTAAAAGTGCAGGCTCTGGAGAGATTAAAGCAGCAGAGAGAGATATCCACAG 2427
2101 AGTGAAGTGTCTATATCTGCTGCTGATGATGAGATGATGAATTAATGATATCCACCA 2160
2428 AGTGAAGTGTCTATATCTGCTGCTGATGATGAGATGATGAATTAATGATATCCACCA 2487
2161 AGACCTGAATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
2488 AGACCTGAATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2547
2281 TCGGAGGCTCATTTTGGGCTTCTGATGCTCCAAATGCTCCCATACCTGATCTCTCCA 2280
2548 TCGGAGGCTCATTTTGGGCTTCTGATGCTCCAAATGCTCCCATACCTGATCTCTCCA 2607
2281 CCTGGCCAAATCACCGAGCTGAAGGAGGAGAAATTCACGGGGGAGCTCATTAATCTGAT 2340
2608 CCTGGCCAAATCACCGAGCTGAAGGAGGAGAAATTCACGGGGGAGCTCATTAATCTGAT 2667
2341 TGGACAGCTCTGGGAGTGAATTAAGACATGAGAAACAGTCAAGATTAATCATTCGAATA 2400
2668 TGGACAGCTCTGGGAGTGAATTAAGACATGAGAAACAGTCAAGATTAATCATTCGAATA 2727
2401 AGTACAGTATTTCTGATCTCAGAGACAGTTCATGATGATGATGATGATGATGATGAT 2460
2728 AGTACAGTATTTCTGATCTCAGAGACAGTTCATGATGATGATGATGATGATGATGAT 2787
2461 GCTCTATCCCAAGAGAGCACTGAGAGAGCTTTTGTGTTAAACAGAAACATT 2520
2788 GCTCTATCCCAAGAGAGCACTGAGAGAGCTTTTGTGTTAAACAGAAACATT 2847
2521 ACTTTGAAATGGACAGATCTTTTCAATGCTATTCAGGCTGTGATTAAGTGAATGCTG 2580
2848 ACTTTGAAATGGACAGATCTTTTCAATGCTATTCAGGCTGTGATTAAGTGAATGCTG 2907
2591 AATTCAGAAATATCAACATTCAGAGAGTATCTTTGTTATTTCTCCACAGACTCGCCA 2640

	DB	2908	AAATCAGAAAATATCCAACTGACGAGACTACTTGTATTATTCCTCCACAGACTCGGCCA	2967
OY	2641	GAGACACCTACTCCGTCGAAGAACGCCTCCTCTTGTTCTTAATATTCATATCAACAGCAC	2700	
DB	2968	GAGACACTAGTCCGATGAAACGCTCTCTTGTCTTAATATTCATATCAACAGCAC	3027	
OY	2701	ATPCTTGCAATTCATATTTAAAAATATATGGAAGTGCATGAGAACCTGAGCTGCA	2760	
DB	3028	ATPCTTGCAATTCATATTTAAAAATATATGGAAGTGCATGAGAACCTGAGCTGCA	3087	
OY	2761	ATPAGCTTAGGGCTGATTTTTGTCTAGATATAATATAATATCATTCACCTT	2812	
DB	3088	ATPAGCTTAGGGCTGATTTTTGTCTAGATATAATATAATATCATTCACCTT	3139	
RESULT	3			
LOCUS	AF127036	2826 bp	mRNA	PRI 10-AUG-1999
DEFINITION	Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)			
ACCESSION	AF127036			
VERSION	AF127036			
KEYWORDS	AF127036.1 GI:4585468			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
TITLE	1 (bases 1 to 2826)			
JOURNAL	Agnel, M., Vermaat, T. and Culouscou, J.-M.			
MEDLINE	Identification of three novel members of the calcium-dependent			
PUBMED	chloride channel (CaCC) family predominantly expressed in the			
REFERENCE	digestive tract and trachea			
AUTHORS	FERS Lett. 455 (3), 295-301 (1999)			
JOURNAL	99364503			
TITLE	10437792			
FEATURES	2 (bases 1 to 2826)			
source	Agnel, M. and Culouscou, J.-M.			
gene	Direct Submission			
CDS	Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des			
	Carrieres, Neuilly-Malmaison 92500, France			
	location=Qualifiers			
	1..2826			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	1..2826			
	/tissue_type="small intestine; colon"			
	5..2749			
	/gene="CaCC1"			
	/note="bovine epithelial chloride channel homolog"			
	/codon_start=-1			
	/product="calcium-activated chloride channel protein 1"			
	/protein_id="AAD25487.1"			
	/db_xref="GI:4585468"			
	/translation="MGPFKSSVFLLIHLLEGALSNGLIONNNGYEIGIYAIDPNP/			
	EDLELQIDKDMYQAOLIVLPFGAKRSEPKRNAILIPMKRYADVYRRPLETKNA			
	DVLVESTPEPPNDDEPTTEOMGNCKEKEERLIHPDPIAGKKLATIGYGRFEMAH			
	LRMGVDFENNDERFYLSNGRIQAVNRCSAIGITNVNKKOGSGCYRKCFNKVTGI			
	YKGECEFLVLRQTERKASIMFAOHVDSIVECEQNHNKSNPNKONOKCMILSTWEVI			
	RDSGEFKTTPMTPOPPNPFTSLQIQIRIVCLVDKSGSMATGNRLNLNOAGLF/L			
	LQYAEISVMGWTFDSAHHQSLEIIQINSGBDRTAKRLPAASAGTSICSGRSRA			
	LYVRKRPYDGSEIVLTGGEEDNTISGCENEWKSQAIIHTVALGPSAOELEELSK			
	MTGLOYYASDOYNNGLIDAFGLMSNCNARSORSLQESKGITTONSOMNNTVIV			
	DSYVGKOTLELITWTPOPOLILMDPSGOKGGEVVDKNRMAYLIQPLAGVGMTRKI			
	SLNASQTLTLTVTSRASNMALPPIVTVTSKTMTKTSFPSPVLVYANI RGASILIRA			
	SVALLIESVNGKVTLLELDNGAGADTKDGYSRFTTYDTNGRSVKVRALGYNVA			
	AARRVYIPDSGSALEYIPGIWENDEIQNNPRPEINKDDVOHKQVCFRTSGSGSFVAS			
	DVENAPLPDPGPQIIDLKAEIHGSGLINTLMTAPGDYDHGAHKYTIIRISLIID			
	LROKENSLQOVNTTALI PREANSEEVFLFPEENTFENGTDHFALIOADVNDLKSEL			
	SNARVSILFIPTQTPEPTSPDESACCPNIHNSITIPGIHLIKIMKWIGCELQLSTA			
	"			
BASE COUNT	875 a 623 c 632 g 696 t			

[illegible]

Db	961	TGGTAACCGGCTCAATGACGTAATCAAGACGGCCAGCTTTCTCTGCTCCAGACAGTTGA	1029
Qy	1041	GCATGGGGTCTCGGGTTGGGATGGTGACATTTGACATGCTGCCCCATGTACAAAGTAAC	1100
Db	1021	GGTGGGGTCTCTGGGTTGGGATGGGACATTTGACATGCTGCCCCATGTACAAAGTAAC	1080
Qy	1101	CATACAGATTAACAGTGGCATGGCATGACAGGACACACTCTGCCAAAAGATTACCTGAGAGC	1160
Db	1081	CATACAGATTAACAGTGGCATGGCATGACAGGACACACTCTGCCAAAAGATTACCTGAGAGC	1140
Qy	1161	TTCCAGGAGGAGCTCCATCTGCAGCGGGCTTCGATGCGCATTTACTGTATTAGGAAGA	1220
Db	1141	TTCCAGGAGGAGCTCCATCTGCAGCGGGCTTCGATGCGCATTTACTGTATTAGGAAGA	1200
Qy	1221	ATATCCAACTGATGGATCTGAAATTGGTGGCTGACAGGATGGGGAAACAAACTATTAAG	1280
Db	1201	ATATCCAACTGATGGATCTGAAATTGGTGGCTGACAGGATGGGGAAACAAACTATTAAG	1260
Qy	1281	TGGGTGCTTTAACGAGGTCAACAAAGTGGTGCATCATCCACACAGTCGCTTTGGGGCC	1340
Db	1261	TGGGTGCTTTAACGAGGTCAACAAAGTGGTGCATCATCCACACAGTCGCTTTGGGGCC	1320
Qy	1341	CTCTGCAGCTCAAGAACTAAGAGAGCTGTCCAAATATACAGAGGTTTACAGACATATGC	1400
Db	1321	CTCTGCAGCTCAAGAACTAAGAGAGCTGTCCAAATATACAGAGGTTTACAGACATATGC	1380
Qy	1401	TTCCAGATCAAGTTCCAGAACATGGCCTCATGTGATGCTTTGGGGCCCTTTCATCAGAAA	1460
Db	1381	TTCCAGATCAAGTTCCAGAACATGGCCTCATGTGATGCTTTGGGGCCCTTTCATCAGAAA	1440
Qy	1461	TGGAGCTGTCTCAGCGGCTCCATCCAGCTTTGAGAGTAAAGGATTTAACCTCCAGAACAG	1520
Db	1441	TGGAGCTGTCTCAGCGGCTCCATCCAGCTTTGAGAGTAAAGGATTTAACCTCCAGAACAG	1500
Qy	1521	CCAGTGGATGAATGGGACAGTGTATCGTGGAGACAGCCGTGGGAAAGAGCACTTGTGTTCT	1580
Db	1501	CCAGTGGATGAATGGGACAGTGTATCGTGGAGACAGCCGTGGGAAAGAGCACTTGTGTTCT	1560
Qy	1581	TATCACCCTGGACAACGACGCTCCCAAATCTTCTGTGGATCCAGTGGACGAAGA	1640
Db	1561	TATCACCCTGGACAACGACGCTCCCAAATCTTCTGTGGATCCAGTGGACGAAGA	1620
Qy	1641	AGGTGGCTTTGTATGGAGACAAAACCAACAAATGGCCTACTCCAAATCCACAGCATATGC	1700
Db	1621	AGGTGGCTTTGTATGGAGACAAAACCAACAAATGGCCTACTCCAAATCCACAGCATATGC	1680
Qy	1701	TAAAGTTGGCACTTGGAAATATACGTGTGCAAAGCAAGCTCAAAACCTTGAACCTGACTGT	1760
Db	1681	TAAAGTTGGCACTTGGAAATATACGTGTGCAAAGCAAGCTCAAAACCTTGAACCTGACTGT	1740
Qy	1761	CACGTCCCGTGGCTCCAAATGCTACCGTGGCTCCAAATTTACAGTGACTTCCAAAACGAACAA	1820
Db	1741	CACGTCCCGTGGCTCCAAATGCTACCGTGGCTCCAAATTTACAGTGACTTCCAAAACGAACAA	1800
Qy	1821	GGACACACACAAATTTCCCGACGCTCTGGTGAATTTATGCAAAATATTCGCCAAGAGCCTC	1880
Db	1801	GGACACACACAAATTTCCCGACGCTCTGGTGAATTTATGCAAAATATTCGCCAAGAGCCTC	1860
Qy	1881	CCCAATTTCAAGGCCAGTGTCAAGCCCTGATTGAATCAGTGAATGAGAAACACGTTAC	1940
Db	1861	CCCAATTTCAAGGCCAGTGTCAAGCCCTGATTGAATCAGTGAATGAGAAACACGTTAC	1920
Qy	1941	CTTGGAAACTACTGATTAATGAGACAGAGTGTGATGCTACTAAAGATAGAGGTGCTACTC	2000
Db	1921	CTTGGAAACTACTGATTAATGAGACAGAGTGTGATGCTACTAAAGATAGAGGTGCTACTC	1980
Qy	2001	AAGGTAATTTCAACGTTATGACACGAATGGTAGATACAGTGTAAAAGTGGGGCTCTGGG	2060
Db	1981	AAGGTAATTTCAACGTTATGACACGAATGGTAGATACAGTGTAAAAGTGGGGCTCTGGG	2040
Qy	2061	AGGAGTTAACGACGACGACGAGAGTGAATACCCACAGATGGAGCACTGTACATATCC	2120
Db	2041	AGGAGTTAACGACGACGACGAGAGTGAATACCCACAGAGTGGAGCACTGTACATATCC	2100

[illegible]

[illegible]

Query Match	Best local similarity	Matches 2256	Conservative	0	Mismatches	531	Indels	9	Gaps	3
892 a	758 c	755 g	674 t							
BASE COUNT										
ORIGIN										
Query Match	67.5%	Score 1898.4	DB 4	Length 3079						
Best local similarity	80.7%	Pred. No. 0								
Matches 2256	Conservative	0	Mismatches	531	Indels	9	Gaps	3		
Query Match	67.5%	Score 1898.4	DB 4	Length 3079						
Best local similarity	80.7%	Pred. No. 0								
Matches 2256	Conservative	0	Mismatches	531	Indels	9	Gaps	3		

Db 957 GATCCAGATTCTGAGACTTTAAGAAAAACCTCCATGACAAACAGCCACCCGCGCC 1016
 QY 909 CACCTTCTCATCTGCTGAGATTGGACAAAGATTGTGTAGTCTCTTGAACAATCTGG 968
 Db 1017 CACCTTCTCATCTGCTGAGATTGGACAAAGATTGTGTAGTCTCTTGAACAATCTGG 1076
 QY 969 AACCATGCGCATGTGTAAACCGCTCAATGCAATCAAGCAGGCCAGCTTCTGCT 1028
 Db 1077 AACCATGCGCATGTGTAAACCGCTCAATGCAATCAAGCAGGCCAGCTTCTGCT 1136
 QY 1029 GCAGACAGTTGACCTGGGGTCTCGGCTTGGGATGTGACATTTGACAGTGTGCCATGT 1088
 Db 1137 GCAGACAGTTGACCTGGGGTCTCGGCTTGGGATGTGACATTTGACAGTGTGCCATGT 1196
 QY 1089 ACAAAAGTAATCTATACAGATAAAGAGTGGCAGTACAGGAGCAGCTCCGCAAAAGATT 1148
 Db 1197 AAAAAAGTAATCTATACAGATAAAGAGTGGCAGTACAGGAGCAGCTTCCGCGACCTT 1256
 QY 1149 ACCTGACAGCTTTCAGAGGAGCTTCATCTGACGCGGCTTCGATCGGCAATTTACTGT 1208
 Db 1257 ACCACAGAGCTTCGAGGAGAGAGCTCATCTGCTCGGGCTTCGCTCCGATTTACTGT 1316
 QY 1209 GATTAGGAAGAAATATTCACACTGATGATGTGAATTTGTCTGTGACGAGTGGGAGAGA 1268
 Db 1317 GATTAGGAAGAAATATTCACACTGATGATGTGAATTTGTCTGTGACGAGTGGGAGAGA 1376
 QY 1269 CAACACTATAGTGGTGTCTTAACGAGTCAAAAGTGTGGCCATCATCCACAGT 1328
 Db 1377 CAACACTATAGTGGTGTCTTAACGAGTCAAAAGTGTGGCCATCATCCACAGT 1436
 QY 1329 CGCTTTGGGGCCCTCTGACAGCTCAAGAACTAGAGAGCTTCCAAAATGACAGAGGTTT 1388
 Db 1437 CGCTTTGGGGCCCTCTGACAGCTCAAGAACTAGAGAGCTTCCGAGATGACAGAGTGT 1496
 QY 1389 ACAGCAATATGCTTCAGATTAAGTTCAGAACAAATGGCTTATGATCTTTGGGGCCCT 1448
 Db 1497 GCAGACATATGCTTCAGATTAAGTTCAGAACAAATGGCTTATGATCTTTGGGGCCCT 1556
 QY 1449 TTTCATCAGAAATGAGACTTCTCTCAGCGCTCCATCCAGCTTGAAGTAAGGATTAAC 1508
 Db 1557 TTTCATCAGAAATGAGACTTCTCTCAGCGCTCCATCCAGCTTGAAGTAAGGATTAAC 1616
 QY 1509 CCTCCAGAACAGCCAGTGTGTAATGGCACAGTATCTGGACAGCAGCCCTGGGAAAGA 1568
 Db 1617 CCTCCAGAACAGCCAGTGTGTAATGGCACAGTATCTGGACAGCAGCTGGGAAAGA 1676
 QY 1569 CACTTTGTTCTTATACCTCGAGCAACGAG----CCTCCCAAAATCCTCTCT- -GGGA 1622
 Db 1677 CACTTTGTTCTTATACCTCGAGCAACGAGGAAATCTTCCCAAAATCTTTTGGGGT 1736
 QY 1623 TCCAGTGGACAGAGCAAGGTGCTTTGTAGTGCACAAAGAACCAAAATGGCTTACT 1682
 Db 1737 TCCAGTGGAGAGAGCAAGGTGCTTTGTAGTGCACAAAGAACCAAAATGGCTTACT 1796
 QY 1683 CCAAAATCCAGGATGCTAAGTGTGGCACTTGAATTAACAGTCTGCAAGCAAGCTTACA 1742
 Db 1797 CCAAGTCCAGGAGCAAGGTGCTGATGTGAGTACAGCTTGCACAAAGCTTACA 1856
 QY 1743 AACCTTGACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1802
 Db 1857 AACCTTGACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1916
 QY 1803 GACTTCCAAAAGCAAGAGACAGCAAGCAATTCCTCCAGCCCTCTGTAGTATTTAGCAAA 1862
 Db 1917 GACTTCCAAAAGCAAGAGACAGCAAGCAATTCCTCCAGCCCTCTGTAGTATTTAGCAAA 1976
 QY 1863 TATTCGCCAAGAGCTCCCAATTTCTCAGGCCAGTGTACAGCCCTGATTTGATCAGT 1922
 Db 1977 GATTCGCAAGAGAGCTCCCAATTTCTCAGGCCAGTGTACAGCCCTGATTTGATCAGT 2036
 QY 1923 GAATGAAAAACAGTACCTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1982
 Db 2037 GAATGAAAAACAGTACCTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2096

QY 1933 GGATGAGGCTGCTACTCAAGATTTTACAACTTATGACAGAGTGTAGATACAGTGT 2042
 Db 2097 GAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2156
 QY 2043 AAAAGTGGGCTCTGAGAGAGTTAAACGACGACGAGAGAGTGTATCCCGACAGAG 2102
 Db 2157 AAAAGTGGGCTCTGAGAGAGTTAAACGACGACGAGAGAGTGTATCCCGACAGAG 2216
 QY 2103 TGGAGCATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2162
 Db 2217 TGGAGCATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2276
 QY 2163 ACCGTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2222
 Db 2277 ACCGTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2336
 QY 2223 GGGAGGCTCATTTTGTGCTTCTGATGTCCCAAAATGCTCCATACCTGATCTCTCCACC 2282
 Db 2337 GGGAGGCTCATTTTGTGCTTCTGATGTCCCAAAATGCTCCATACCTGATCTCTCCACC 2396
 QY 2283 TGGCAATATACCGACTGAAAGGCGGAATTTACAGGGGAGTGTCTATATCTGACTTG 2342
 Db 2397 CTGTAAGATACCGACTGAAAGGCGGATTCGAAGGCGGAGACAAACCTCATCAACCTGAGCTG 2456
 QY 2343 GACAGCTCTGGGGATGATTAATGACATGGAACGCTCACAAAGTATTCATTCGATTAAG 2402
 Db 2457 GACAGCTCTGGGGATGATTAATGACATGGAACGCTCACAAAGTATTCATTCGATTAAG 2516
 QY 2403 TACAAGTATCTTCGATCTGACAGAGCAAGTTCATTAATGATCTCTCAAGTATGATCTGCTG 2462
 Db 2517 TACAAGTATCTTCGATCTGACAGAGCAAGTTCATTAATGATCTCTCAAGTATGATCTGCTG 2576
 QY 2463 TCTCATCCCAAGAGCAAGCTGAGAGAGCTTTGTTGTTAAACAGAGGACATCC 2522
 Db 2577 TCTCATCCCAAGAGCAAGCTGAGAGAGCTTTGTTGTTAAACAGAGGACATCC 2636
 QY 2523 TTTTGAATATGACAGATCTTTTCAATGCTATTCAGGCTGTGATTAAGTGTGATCGATCGAA 2582
 Db 2637 TTTTGAATATGACAGATCTTTTCAATGCTATTCAGGCTGTGATTAAGTGTGATCGATCGAA 2696
 QY 2583 ATCAGAAATATCCAACTTGCAGAGATCTTTGTTATTCCTCCACAGACTCCCGAGA 2642
 Db 2697 GTCAAGAAATATCCAACTTGCAGAGATCTTTGTTATTCCTCCCGAGAGCTCCCGAGA 2756
 QY 2643 GACACCT---AGTCTATGAAGCTGCTCTCTGCTCCAAATTTGATATTCACACAGCAG 2699
 Db 2757 GACTCTCCGAGACACCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2816
 QY 2710 CATTCTCGCATTCACATTTTAAATTTATGTGAATGATGATGATGATGATGATGATGATGATGAT 2759
 Db 2817 CATTCTCGCATTCACATTTTAAATTTATGTGAATGATGATGATGATGATGATGATGATGATGAT 2876
 QY 2750 AATAGCCTAGGCTGAATTTTGTCAATTAATAA 2795
 Db 2817 GATAGCCTAGGCTGAATTTTGTCAATTAATAA 2912
 RESULT 6
 AB017156 2937 bp mRNA ROD 10-NOV-1999
 DEFINITION Mus musculus gob-5 mRNA, complete cds.
 ACCESSION AB017156
 VERSION AB017156.1 GI:3721911
 KEYWORDS GOB-5;
 SOURCE Mus musculus
 T. ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (stiles)
 Komiyama, T., Tanigawa, Y. and Hirohashi, S.
 Cloning and identification of the gene gob-5, which is expressed in
 intestinal goblet cells in mice

JOURNAL Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)
MEDLINE 99160866
REFERENCE 2 (bases 1 to 2937)
AUTHORS Komiyama, T., Tanigawa, Y., and Hirohashi, S.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1998) to the DDBJ/EMBL/GenBank database

Db 1455 GCGGGGATCGTCACACTCATCCAGCTGAGAGCAGGGGAGTAATCTCCAGAAATAC 1514
 QY 1522 CAGTGATGATGAGGACAGTGTGATGAGAGACAGCGTGGGAAAGACACTTTGTTCTT 1581
 Db 1515 CAATGATGATGAGGACAGTGTGATGAGAGACAGCGTGGGAAAGACACTTTGTTCTT 1574
 QY 1582 ATACCTGAGCAACGACAGCTCCCAATCTCTCTGAGATCCAGTGGACAGAAACAA 1641
 Db 1575 ATACCTGAGCAACGACAGCTCCCAATCTCTCTGAGATCCAGTGGACAGAAACAA 1634
 QY 1642 GGTGGCTTTGTAGTGGACAAACAAACCAAAATGGCTACCTCAATCCAGGCAATGCT 1701
 Db 1635 AATGCTTTTATCTAGTGGACAAACCAAAATGGCTACCTCAATCCAGGCAATGCT 1694
 QY 1702 AAGGTTGACCTTGGAAATACAGTGTGAGAGGCTCAAAACCTTGGACCTGACCTGTC 1761
 Db 1695 AAGGTTGACCTTGGAAATACAGTGTGAGAGGCTCAAAACCTTGGACCTGACCTGTC 1754
 QY 1762 AGCTCCCGTGCCTCAATGCTACCTGCTCCCAATTCAGTACCTTCCAAACGAAACAG 1821
 Db 1755 ACCTCCCGTGCCTCAATGCTACCTGCTCCCAATTCAGTACCTTCCAAACGAAACAG 1814
 QY 1822 GACACGACAAATTCCTCCAGCCTCTGAGTATTCAGCAATATTCGCCAAGAGCCTCC 1881
 Db 1815 AACACAGGAAATTCCTCCAGCCTCTGAGTATTCAGCAATATTCGCCAAGAGCCTCC 1874
 QY 1882 CCAATTCAGGAGCCAGTGTGAGAGGCTCAAAACCTTGGACCTGACCTGTC 1941
 Db 1875 CCTATTCAGGAGCCAGTGTGAGAGGCTCAAAACCTTGGACCTGACCTGTC 1934
 QY 1942 TTGGAATCTGATGATGAGAGGCTGTGATGATGATGATGATGATGATGATGATGAT 2001
 Db 1935 CAGGAATCTGATGATGAGAGGCTGTGATGATGATGATGATGATGATGATGATGAT 1994
 QY 2002 AGGATTTTCAACAATTCATACAGAGTGTGATGATGATGATGATGATGATGATGAT 2061
 Db 1995 AGGATTTTCAACAATTCATACAGAGTGTGATGATGATGATGATGATGATGATGAT 2054
 QY 2062 GGAGTTAACGACAGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2121
 Db 2055 GGAGTTAACGACAGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2114
 QY 2122 GGCTGATTTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2181
 Db 2115 GGCTGATTTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2169
 QY 2182 GATGTTCAACAAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2241
 Db 2170 -ATGTTCAACAAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2228
 QY 2242 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298
 Db 2229 ACCAATGTCCTCCGACAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2288
 QY 2299 CTTGAAGGAGGAAATTCACAGGAGGAGTGTGATGATGATGATGATGATGATGAT 2358
 Db 2289 CTTGAAGGAGGAAATTCACAGGAGGAGTGTGATGATGATGATGATGATGATGAT 2348
 QY 2359 GATTTAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2418
 Db 2349 GATTTAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2408
 QY 2419 CTTGAGAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2478
 Db 2409 CTTGAGAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
 QY 2479 GCGAATCTGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2538
 Db 2469 GCGAATCTGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2528
 QY 2539 GATCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2598
 Db 2529 GATCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2588

QY 2599 ATTCAGAGATATCTTTGTTATCTCCACAGACTCCGACAGACACCTAGTCTGAT 2658
 Db 2569 ATTGCACAGGCTGTCTGTGTTATCTCCGCTCAG-----GAGCCGCCATTTCCGGA 2639
 QY 2659 GAAAGCTGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2718
 Db 2640 GACTCAAGCTCCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2699
 QY 2719 TTTAAATTTATGAGAGTGTGATGAGAGACCTGATGATGATGATGATGATGATGAT 2778
 Db 2700 CTGAAGATTAATGAGAGTGTGATGAGAGACCTGATGATGATGATGATGATGATGAT 2759
 QY 2779 TTTGTCAGATTAATTA 2795
 Db 2700 TTTGTCAGATTAATTA 2776
 RESULT 7
 AF127035 3204 bp mRNA PRI 11-AUG-1999
 LOCUS Homo sapiens calcium-activated chloride channel protein 2 (CaCC2)
 DEFINITION mrna, complete cds.
 ACCESSION AF127035
 VERSION AF127035.1 GI:5726288
 KEYWORDS
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 3204)
 Agnel, M., Vernet, T., and Culouscou, J.M.
 Identification of three novel members of the calcium-dependent
 chloride channel (CaCC) family predominantly expressed in the
 digestive tract and trachea
 FEBS Lett. 455 (3), 295-301 (1999)
 JOURNAL MEDLINE 99364503
 PUBMED 10437792
 REFERENCE 2 (bases 1 to 3204)
 Agnel, M. and Culouscou, J.-M.
 Direct Submission
 Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des
 Carrières, Ruell-Malmaison 92500, France
 JOURNAL Location/Qualifiers
 FEATURES
 source
 gene
 cds
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="colon"
 1. 3204
 /gene="CaCC2"
 29. 2782
 /gene="CaCC2"
 /note="bovine epithelial chloride channel homolog"
 /codon_start=1
 /product="calcium-activated chloride channel protein 2"
 /protein_id="AAD48398.1"
 /db_xref="GI:5726288"
 /translation="MGLPRGFLVLLICLLHOSNPSFILTNNNGEDIVIVDPSP
 DEFTIIOIDMTASTYLFEATKFFKNSILPEMKKEPOYKRRKHEHKAH
 VIVAPETLGRREPKKTEGEKEVYHIFPDLLGRKMEYGPGLFVHEMAH
 RMGVEDEVNDQPFYRAKSKLEATKCSAGISGRNRYVCGGSCSRACRIDSTKL
 YKDCOFDPDKOTERAKSIMPQSIDSVAFECHEHNOEAPLOIKNFRSTYKL
 SNSEDKNTIPVPTPPPPVPSILKIRORIVCLVIDKSGSMGKDLNMMOAKHFL
 LQTVENGSGVAVHDSPTATYVKKLQIKSSDERNTLMLGPTPLPGSTISGITYA
 FOYIGLHSLDSEVLLITGDEHNTASSCTIDVKOSGAIHVHTALGRADEVITMS
 KITGSHFVSDAONNGLIDAFGALTSGNTDLSQSLQESGLTNSNAMNDTVI
 IDSTVGKDFEFLITVNSLSPSISIMPSGTLNEMKNDVNSPSPMTVVAEILQGYV
 VYLOAKANBEFLITVNSLSPSISIMPSGTLNEMKNDVNSPSPMTVVAEILQGYV
 LGANTYAFLESNGHTEVLELDNGAGASPKMDGYVSTFAYATENGYSIKRVAG
 GANTALIKRPLPLNRAVITPCVYVNEIENANPREPIDEDTDTOTTEDSRTASGAFV
 VSOVSLPLPDPTPSQITDLDVTHEDTILTWAPGNFDGKQVRIITRISATIL
 DLRSDFDDLDQVTTDLSPKENSKESEFAKPEENISEENATHTFIAIKSIDSNLTSK
 VSNIAQVTLFIQANDPIDPTPTPDKSHNSGVNISTVLVSGSVIVNFTLS

Db 756 GTTATGCAAGATGATCTGTGTTGATTTGTAACGAAAAAACCATATACAGA 815
 Oy 798 AGCTCCAAACAGCAAAATATGCAATCTCCAGACATGGAGATGATCCGTGA 857
 Db 816 AGCTCCAAAGCTCAAAACATTAAGTGCATTTTAAAGTACATGGAGGTGATTGCA 875
 Oy 858 TTCTGAGGACTTTAAGAAAAACACTCTATGACACACAGCCACCAATCCACTTTC 917
 Db 876 TTCTGAGGATTTTAAAAACCATACCATGACACACACCTCTCCACCTCTTCTC 935
 Oy 918 ATTGCTGAGATGACAAAAGATTTGTGTTTAAAGTCTTGACAAATCTGGAAGCATG 977
 Db 936 ATTGCTGAGATCTGCAAAAGATTTGTGTTTAAAGTCTTGGAAGCATG 995
 Oy 978 GACTGTAAACCCGCTCAATGATGATGATGATGATGATGATGATGATGATGATG 1037
 Db 996 GGGTAAAGAGACGCTCAATGATGATGATGATGATGATGATGATGATGATGATG 1055
 Oy 1038 TGAGCTGGGCTCTGGGTTGGATGATGATGATGATGATGATGATGATGATGATG 1097
 Db 1056 TGAAGTGGATCTGGGTTGGATGATGATGATGATGATGATGATGATGATGATG 1115
 Oy 1098 ACTCATACAGATTAACAGTGGATGATGATGATGATGATGATGATGATGATGATG 1157
 Db 1116 GCTAAATCAAAATTAAGAGATGATGATGATGATGATGATGATGATGATGATG 1175
 Oy 1158 AGCTTCAGAGAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1217
 Db 1176 TCCTCTGGAGAGACCTGATGATGATGATGATGATGATGATGATGATGATGATG 1235
 Oy 1218 GAATAT---CAACTGATGATGATGATGATGATGATGATGATGATGATGATG 1274
 Db 1236 GCTACATCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1295
 Oy 1275 TATATGAGGCTCTTAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1334
 Db 1296 TGCAGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1355
 Oy 1335 GGGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1394
 Db 1356 GGGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1415
 Oy 1395 ATATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1454
 Db 1416 TTATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1475
 Oy 1455 AGGAATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1514
 Db 1476 AGGAATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1535
 Oy 1515 GAACAGCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1574
 Db 1536 TAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1595
 Oy 1575 GTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1634
 Db 1596 CTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1655
 Oy 1635 GAAGCAAGTGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1694
 Db 1656 AATATGAAAAATTTACAGATGATGATGATGATGATGATGATGATGATGATGATG 1715
 Oy 1695 CATTCCTAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1748
 Db 1716 AACTCAAGAGTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1775
 Oy 1749 GACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1808
 Db 1776 AACTTATCAAGTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1835
 Oy 1809 CAAAAGCAAGAGACACAGCAATTTCCAGGCTCTGATGATGATGATGATGATGATG 1868

Db 1836 TAAATGATTAAGAGATTAAGATTTCCCAAGCCCAATGATGTTTACGAAATTTCT 1895
 Oy 1859 CCAAGAGAGCTCCCAATTTCTCAGGCGCCAGTGTACAGCCCTGATTTGAATGATG 1928
 Db 1896 ACAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1955
 Oy 1929 AAAACAGTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1988
 Db 1956 ACATACAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2015
 Oy 1989 CGGTCTATCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2048
 Db 2016 TGAAGTCTATCTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2075
 Oy 2049 GGGGCTCTGAGAGATTAAGCAAGCCAGAGAGATGATGATGATGATGATGATGATG 2108
 Db 2076 TGGGCTCTGAGAGATTAAGCAAGCCAGAGAGATGATGATGATGATGATGATGATG 2135
 Oy 2109 ACTGTATACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2168
 Db 2136 CCGTATACATACAGAGCTGGGATGATGATGATGATGATGATGATGATGATGATGATG 2195
 Oy 2169 AATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2228
 Db 2196 AATGAT---GAGATATCTAGAGATGATGATGATGATGATGATGATGATGATGATG 2252
 Oy 2229 CTTATTTGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2288
 Db 2253 TGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2312
 Oy 2289 AATCAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2348
 Db 2313 AATCAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2369
 Oy 2349 TCCCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2408
 Db 2370 ACCAGAGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2429
 Oy 2439 TATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2468
 Db 2470 TATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2489
 Oy 2489 CCCAAGAGAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2528
 Db 2520 ACCAAGAGAGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2549
 Oy 2529 AAATGCAAGATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2588
 Db 2550 AAATGCAAGATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2609
 Oy 2599 AATATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2645
 Db 2610 AGATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2669
 Oy 2616 ACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2683
 Db 2670 TCTTACACCTACTCTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 2707

RESULT: 9
 AX092338
 Locus 3265 bp DNA PAT 21-MAR-2001
 DEFINITION Sequence 69 from Patent W00116318.
 AX092338
 VERSION 1
 KEYWORDS: human
 SOURCE Human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3265)
 Patton,D.L., Fliviaroff,E., Gerritsen,M.E., Goddard,A.,
 Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and

Db 1938 ACATACAGAGAGTTTGGAACTTTGGATAATGTCAGAGCGCTGATTTTCAAGAATGA 1997
 Oy 1989 CGGTCTACTACAGGTATTTTACAACTTATGACAGSAGTGTACATGAGTAAAGT 2048
 Db 1998 TGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAATGGCAATATAGCTTAAAGT 2057
 Oy 2049 GCGGCGCTCTGGAGAGAGTTAAACGACAGAGAGAGTGAATCCCGACAGAGTGCAGC 2108
 Db 2058 TCGGCGCTCATGAGAGAGCAACAGCTGAGGCTAAATATAGCGCTCCACGTAATAGAGC 2117
 Oy 2109 ACTGTACATCTGCTGCTGATTTGAGATGATGAATACATGGAATCCACCAAGCTTGA 2168
 Db 2118 CGCGTACATACAGGCTGGGTAGTGAACGGGAAATTTGAAGCAACCCGCCAAGACCTGA 2177
 Oy 2169 AATTAATAGATGATGTTTCAACAGCAAGAGTGTGTTTACGACATACATCTCGGAGG 2228
 Db 2178 AATTTAT---GAGGATCTACAGACCACTTGGAGATTTTACCGACAGCATCCGAGG 2234
 Oy 2229 CTGATTTGTGCTTCTGATGTCCTCAATGCTCCATACCTGATCTCTTCCACCTGGCCA 2288
 Db 2235 TGCATTTTGTGATTCACAAAGTCCCAAGCTTCCCTTGGCTGACCAATACCCCAAGTCA 2294
 Oy 2289 AATCCACCACTGAGAGCGGGAATTCACGGGGGCACTCTATTATGCTGATGAGAC 2348
 Db 2295 AATCCACCACTGATGACAGAGTTCATGAGG---ATAAGATTATCTTACATGAGACAGC 2351
 Oy 2349 TCTCTGGATGATTTATGACATGAGAGCTGACAAATATATCATTCGATTAAGTACAG 2408
 Db 2352 ACCAGGATATATTTGATGTTGGAAGGTTCAAGCTTATATCATTAAGATTAAGTGCAG 2411
 Oy 2409 TATTTCTGATCTCAGACAGCAAGTTCAATGATCTCTCAAGTAACTACTGCTCAT 2468
 Db 2412 TATTTCTGATCTCAGACAGCAAGTTGATGATGCTCTCTCAAGTAACTACTGATCTGTC 2471
 Oy 2469 CCCAAGGAGGCACTCTGAGAGAGTCTTTTGTGTTTAAACCAAGAACTTACTTTGA 2528
 Db 2472 ACCAAGGAGGCACTCTGAGAGAGGCTTTGATTTAAACCAAGAAATATCTCAGAGA 2531
 Oy 2529 AATGCGACAGATCTTTTCAATGCTATTCAGGCTGTGATTAAGTGCATGCTGAATGAGA 2588
 Db 2532 AATGCGACCAACATATTTATTTGATTAAGATTAAGTAAAGCAATTTGACATCAAA 2591
 Oy 2589 AATATCCACATGTCAGAGATCTTTGTTTATTCCTCCACAGACCTCCGACAGAGA---C 2645
 Db 2592 AGTATCCACATGTCAGAGATCTTTGTTTATTCCTCCACAGAAATCTGATCATTTGA 2651
 Oy 2646 ACCTAGTCTGATGAAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2683
 Db 2652 TCTACACTACT 2689
 RESULT 10
 BTU36445 2984 bp mRNA MAM 09-FEB-1996
 LOCUS BTU36445 2984 bp mRNA MAM 09-FEB-1996
 DEFINITION Bos taurus calcium-activated chloride channel mRNA, complete cds.
 ACCESSION U36445
 VERSION U36445.1 GI:1184065
 KEYWORDS
 SOURCE
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 2984)
 Cunningham, S.A., Awayda, M.S., Buhien, J.K., Ismailov, I.I.,
 Arrate, M.P., Berdiev, B.K., Benos, D.J., and Fuller, C.M.
 Cloning of an epithelial chloride channel from bovine trachea
 J. Biol. Chem. 270 (52), 31016-31026 (1995)
 MEDLINE 96125078
 REFERENCE 2 (bases 1 to 2984)
 Benos, D.J.
 AUTHORS
 TITLE Direct Submission

JOURNAL Submitted (15-SEP-1995) Dale J. Benos, Physiology & Biophysics,
 University of Alabama at Birmingham, BHSB 706, Birmingham, AL
 35294, USA
 FEATURE Location/Qualifiers
 1..2984
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /cell_type="epithelial cells"
 /tissue_type="trachea"
 19..2730
 /note="epithelial chloride channel"
 /codon_start=1
 /product="calcium-activated chloride channel"
 /protein_id="AAC48511.1"
 /db_xref="GI:1184065"
 /translation="MVRPLVYLFLTHLPKSSKSVNLLNNGYDVIATLNSVE
 DEKLQIKRENVLEASTYLEHATRRRYFRNVLILIPMTKSEYLPPOESDME
 VIVANPLKHDDPYTLQYRCGERGYIHTPNPLILNMLPIYSGRAFVHEMAH
 RMGJFEDYNDGPYISRNMTIATRCSTHTTGNNVLYKCGGSCIRPDRDSOTL
 YEAKCTFLPEKSORARSLMEPMOSLHSTVECTECPNVEAPNLONKMGKSWDVI
 MNSDPQNTSPTEMPNPPTPTESILSKRYVCLYDKSGMSSEDRLRMNAEL
 FLQIIEKSLVGVTFDSVAEINNLTKITLDVNYENTITANLPOEANGGISICRGLK
 AGFOAIIQSOSTISSEILITLDEDEHISCLEEVKQSEVLIHTLALGSAELEET
 LMDTGHRREYANKDINGLNNASRISRSQSTIQOTLOESKALATERKKNWCTVP
 LVDYIGNDTFEVVYTIKPEILLQDPRKKYKSDPEKDLNHSARLIPVAGATG
 TWYSLLNNHASPQILFVYVTRARSPPTPTATAMSONTARHPSPVIVAGVSOG
 FLPLVGINVTAIIETEDGHQVTELMNGAGADVKKNDGYSRYFDYRNGRSILYV
 HAERNNRARSLSIQPONKALIXGYIENKILINPRPREYKDLAAEIEDSRILTS
 GGSFTVGAAPGNHPSVLPKRIIDLEAKREDITQLSWPAVAVLKGANSTIIRI
 SKSFLDLOKPDNATLVNTSLKPKREGSDENEFKPEPRIENTGNVFLAVOAINBA
 NLTSVENINQALIKFIPMPEDSVAPLOTKISALINLALFALAMLIISIV"
 BASE COUNT 1027 a 583 c 561 g 813 t
 ORIGIN
 Query Match 33.2%; Score 933.4; DB 4; Length 2984;
 Best Local Similarity 61.9%; Pred. No. 1,7e-248;
 Matches 1635; Conservative 0; Mismatches 966; Indels 42; Gaps 8;
 Db 13 AGATGTACAGCAATGGGGCCATTAAAGTCTGTGTCATCTTGTATCTTCTACCTTCTA 72
 Oy 7 ACATGCGCAAAATAGTCTCGTCTGCTGCTCATCTGTTCTTCACTTCCACCTCTCTG 66
 Db 73 GAAGGGGCGCTGAGTAATCTACATTCAGCTGAACAACAATGCTATGAGGATGTC 132
 Oy 57 CCTGG---AATGAAAGTGAATGTAATTTGATTAACATGATTAAGTGGCATGTC 123
 Db 153 GTTGCAATCGACCCCAATGTCGCCAGAGATGAACACTCATTCATCAACAATTAAGACAT 192
 Oy 124 AATGCAATTAACCTAGTGTGCCAGAAAGATGAACATCAATCAAAACATTAAGAAATG 183
 Db 193 GTGACCCGAGCATCTGCTATCTGTGTAAGCTCTGAGAAAGCATTTTATTTAAAT 252
 Oy 184 GTTAAGTCAAGCTTCTTACTTCTGTTTCATGCCACCAACGAAGATTATTTAGAGAT 243
 Db 293 GTTGCCATTTTGAATTCCTGAACAATGAAGCAAGGCTGACTATGTCAGACCAAACTT 312
 Oy 284 GTAAGCATTTTAATTCGAATGACGGAAGTCAAAATCTGATGTTACTTAATGCCAAACAA 303
 Db 313 GAGACCTTAACAAAATGCTGATGTTCTGTTGCTGAGTCTACTCTCCAGTAATGATGAA 372
 Oy 314 GAATCATATGACCGACGAGAGATGATAGTCTATCTCTTCAAAAACATGAGAGAT 363
 Db 333 CCCTACACTGACGAGTGGGCACTGGGAGAGAAAGGTGAAGAGTCCACTACTACT 432
 Oy 334 CCCTATACACTCAATATGAGAGATGAGAGAAAGGACATATATATACATTTTATCTCA 423
 Db 433 GATTTCATTGACAGAAAAGTACTGTAATATGACACACAGCTAGGCAATTTGTCAT 492
 Oy 434 AACTTCTGTGTAATTAATTTGCTATCTATGTGGTCCGAGGCAAGCATTTGTCTAC 483
 Db 43 GAGTGGCTCATCTACGATGGGAGTATTTGACAGATCAATTAATGATGAGAAATTTGAC 552

Db 484 GAGTGGGCCAATCTCCGATGGGGAATATTGATGATTAACGGGACCAAGCCATCTAT 543
 Oy 553 TTATCCAAATG---AAGATACAGCAGTAAGATGTTTCAGCAGGATATTACTGTACAAAT 609
 Db 544 ATCTCAGAGAGAACACTATTTGAGCAACAGATGTTCAACACTATTTACTGTACTAAT 603
 Oy 610 GTAGTAAGAGATGTCAGGAGGACGCTGTTACACCAAAAAGATGACATTTCAATTAAGTA 669
 Db 604 GTGATTTGTCAAATGCGAGGAGGACCTGTATTAACAGGCGATGACAGCTGACTCAG 663
 Oy 670 ACAGAGCTCATGAAAAAGAGTGTGTTGTTCTCCAAATCCCGCAGAGGAGGAAAGCT 729
 Db 664 ACAGGCGCTGATGAGCAAAATGTACATTCATCCAGAAAATCCAGACTGCAAGGAG 723
 Oy 730 TCTATATGTTTTCACACACATGTTGATTTCTATGATTTGAATTTCTGACAGAACAAACAC 789
 Db 724 TCCATCATGTTTATGCAAGGTCTCCATTTCTGATGCAATTTTGTACAGAAAAACACAC 783
 Oy 790 AACAAAGAGCTCCAAACAGCAAAATCAAAAATGCAATCTCCGAGACATGGAAGTG 849
 Db 784 AATGTGAAGCTCCAAACCTACAAACAAAATGTGCAATGCAAAAAGCACATGGATGTA 843
 Oy 850 ATCCGATTTCTGAGACCTTTAGAAAACCCATCTATGCA-----ACACAGCCACCA 903
 Db 844 ATCATGAACTTACTACTACTTTCAAAATACATCTCCCATGACAGAAATGAATCCACGACT 903
 Oy 904 AATCCACCTTCTCATTTGCTGACAGATTGACAAAGAAATGTGTTAGCTTGTGACAA 963
 Db 904 CAACCTACATTTTTCATTTGCTGCAAGTCCAAACACGAGTACTTTGTTGGTACTGTAATA 963
 Oy 964 TCTGAGACATGCGAGCTGTAACCGCTCATCTGACTGATGCAAGCAGGCGACTTTTC 1023
 Db 964 TCTGAGAGATGCTTCAGAGAGATCTCTCTTTCGATGATGAATGAGAGAGCAATTTATTC 1023
 Oy 1024 CTGCTGCAAGACAGTGTAGCTGGGGTCTCGGTGGATGGATGCTGATTTGACATGCTGCC 1083
 Db 1024 TTGATTTCAATTTATTAAGAGATCTGTTGGATGGATGTTACATTTGACAGTGGTCT 1083
 Oy 1084 CATGTACAAAGTACATCTATACAGATTAACAGTGGCAGTACAGGACACACTGCGCAAA 1143
 Db 1084 GAAATCGGAAATATTAACAAATAATACATGATTAATTTTATGAAAATATTCACATGCA 1143
 Oy 1144 AGATTACCTGACAGCTTTCAGAGGAGAGCTCATCTGACGCGGCTTGCATGGCATTT 1203
 Db 1144 AATCTCCTTCACAGAGCTTAATGTTGGAACTTCAATTTGTAGGGGCTCAAGACAGATTT 1203
 Oy 1204 ACTGTATTTAGGAAGA---AATATCCATGATGATCTCAAAATGTGCTGCTGACGAGAT 1260
 Db 1204 CAGGCAATTTATTCAGATCAGATCAGAGACCTTGTGTTGTAATTCATTAATCCGAT 1263
 Oy 1261 GGGGAGACAAACACTTAATAGTGGTCTTAAAGAGTCAACAAAGTGGTGCATCATC 1320
 Db 1264 GGGGAGATTAATTAATTAACACTCATGATTTGAGAGGTAAACAAAGTGGTGTGATCAT 1323
 Oy 1321 CACACAGTCTTGGGGCCCTCTGACGCTCAAGAACTAGAGAGCTGTCCAAATGACA 1380
 Db 1324 CACACATTTGCTGGGACCTTCTGCTGCCAAAGAACTGAGACACTGAGATATGACA 1383
 Oy 1381 GAGGTTTACAGACATATGCTTCAATCAAGTTCAGAACAAATGGCTCATTTGCTTTT 1440
 Db 1384 GGGGAGACATCGTTTATGCAATTAAGACATA-----AATGGCTTACTTAATGCTTC 1437
 Oy 1441 GGGGCGCTTTTCATCAGAAATAGAGCTGTCTCAGCGCTCCATCCAGCTTGAAGATAG 1500
 Db 1438 AGTAAATTTTCACTAGAGTGGCAGATATCATTACAGACAGCTATTCAGTTGGAAGCAA 1497
 Oy 1501 GGAATTAACCTCCAGAAACGCCAGTGAATGACACAGTATCTGGAACAGACCGTG 1560
 Db 1498 GCTTGGCATTAACAGAAAAAGAAATGGTAAATGCTACGCTGTGATGATACAAAT 1557
 Oy 1561 GGAAGGACATTTGTTTCTTATCCTGAGACAGCAGCCTCCCAAAATCTCTCTG 1620
 Db 1558 GGAATGACATTTCTTTGTTGTCATGACATGACAAATTAAGGCAAGAAATTTCTTCAG 1617

Oy 1621 GATCCAGTGGAGAGAACA-----AGTGGCTTTGTAGTGAACAAAACACCAAA 1671
 Db 1618 GATCCAAAGAGAAATTAATAAACCCTCAGATTTCAAGAAAGATAGCAATATTTATCAT 1677
 Oy 1672 ATGGCTTACCTCCAAATCCAGGCAATGTCATAGTGGTGGACCTTGGAAATACAGTCT--- 1727
 Db 1678 TCTGCTGCTTGGAAATCTGATTTGATGAGAGCAGAGTACTGGACTTGCAGCTTCTTA 1737
 Oy 1728 -----GCAGCAGCTCCAAACCTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1782
 Db 1738 AATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1797
 Oy 1793 ACCCTGCTCCAAATTCAGTACTTCCAAAACGAAAGACACCAACAAATTTCCAGC 1842
 Db 1798 ACTACACCCCGAATGCAACAGCTCACATGATGCAAAATACAGCACATACCTTAGC 1857
 Oy 1843 CCTGCTGATTTATGCAATATTCGCCAAGAGACCTCCCAATTTCTCAGGCGCATGCTC 1902
 Db 1858 CCAAGTATGTTTATGCAAGTCAAGGCTTTTGGCTGTGATGGAATCAATGTA 1917
 Oy 1903 ACAGCCCTGATGATGCAAGTGAATGSAAAAACGTTACCTTGAACACTGATGATGTA 1962
 Db 1918 ACTGCTATTTAGAACTGAAGATGGCCATCAAGTACACTGAGCTTGGGACAAATGCT 1977
 Oy 1932 GCAGTGTCTGATGCTACTAAGATGACGCTGCTACTCAAGTATTTTCAACACTTATGAC 2022
 Db 1978 GCAGTGTCTGATGCTGCAAGAAATGATGATGATGATGATGATGATGATGATGATGAT 2037
 Oy 2023 ACGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2082
 Db 2098 GGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2097
 Oy 2093 AGAGTGTATCCAGCAGATGAGCAGCTGATGATGATGATGATGATGATGATGATGATGAT 2142
 Db 2098 AGTTTAAGACACACGAAACAAAGCCCTGATATACAGCTGATATGATGATGATGATGAT 2157
 Oy 2143 ATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2202
 Db 2158 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2217
 Oy 2203 TGTTCAGCAAAATCTGCGGAGCTCATTTGTGCTGATGATGATGATGATGATGATGATGAT 2262
 Db 2218 GACTTTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2277
 Oy 2263 ATACCTGATGCTTCCACCTGCGCAATGACGACCTGAGGAGGGAATTCACGGGGGC 2322
 Db 2278 CACCTTGTGCTGCTCCACCAATTAATTAATGATGATGATGATGATGATGATGATGATGAT 2335
 Oy 2323 AGTCTATTAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2382
 Db 2386 ATACATTTCACTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2394
 Oy 2393 AAGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2442
 Db 2395 AGCTACATTTAAGAAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2454
 Oy 2443 CTTCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2502
 Db 2455 ACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2514
 Oy 2503 TTTAAACCAAAACATTTATTTGAATGAGCACAATCTTTTTCATTTGATTTAGCT 2562
 Db 2515 TTTAAACCAAAACATTTATTAATGAAATGAGCACAATCTTTATTTGATGATGATGATGATGAT 2574
 Oy 2563 GTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2622
 Db 2575 ATCAATGAAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2634
 Oy 2623 CCT 2625
 Db 2635 CCT 2637

RESULT 11
AF001261 3317 bp mRNA MAM 18-NOV-1997
LOCUS Bos taurus clone 1 endothelial adhesion molecule Lu-ECAM-1 mRNA,
DEFINITION complete cds.
ACCESSION AF001261
VERSION AF001261.1 GI:2623762
KEYWORDS
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 3317)
Eldle, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
Goodwin, A., and Pauli, B.U.
TITLE Cloning and characterization of Lu-ECAM-1 suggest it is an
endothelial chloride channel
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3317)
Eldle, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
Goodwin, A., and Pauli, B.U.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (25-APR-1997) Vet. Pathology, Cornell U., WMC C4-146,
Ithaca, NY 14853, USA
FEATURES
source
Location/Qualifiers
1..3317
/organism="Bos taurus"
/db_xref="taxon:9913"
/cell_type="aortic endothelial cells"
/clone="1"
/tissue_type="lung"
63..2780
/note="endothelial adhesion molecule; chloride channel"
/codon_start=1
/product="Lu-ECAM-1"
/protein_id="AAB86529.1"
/db_xref="GI:2623763"
/translation="MVLNVILFLTLHLPGKSSMVNLINNGYDGLVIAINPSVE
DEKLINIKEMVETASTYLFHATKRYFRVAVSLIPMTKSEVEFLPKESDAD
VIVANPLYKGGDDPYTLOYGRGEGKGYIHETPNLILNNHITGSGRVHVMH
RMGIEDVNDOPPYISRKNTIEATRCSTHITGINVFKKPGSCITISCRISQIG
LYEAKCTFLPKRSOTAKESIMFMSLSHVECEKTHNTAPMLONKMKGKSTWY
IMNSVDQNTSPMTENPPIHPTSLKSKQVVCVLNDSGSASDRLEQMAQAE
LYLTOVEKSLVGMVTFDSVAEIONHLRTIDNVOKITAKIPOVANGSTICREL
KAGFOALIHSDOSTSGSEIILLDGEENEINSCEDYKRSQALIHITLIGSAKKE
TLSNMGGYREFANKDITGLNAPSRISRSQSTLOALIOLESKALITGKRVNGV
PVSTVNDPFEVVTWITOKDEIVLOPKCKKYTSPEKELMIRARQIPDAET
GTPYTSILNNHASSOMLTIVYTTTRASPITIPVIAHMSOHTAHIRSPMLVAQVSO
GFLPVLGISYAIILETDGHQVITLMDNGRDTVANDGISYFTDYGNGRSLK
VHAQRNNTARLNROPNKVLYVGVVENKILINPRPEVKDILAKAKEDSSRLT
SGSFTVSGAPPGNHPSPKSTIDLEAKFKDYIQLSWTAIGNVLDGKAKASYII
RISKSPMDRODFDNATLVNLSNLIPEKAGSEKREKPEHFRVENGKIFYISVOAIN
EANLISVSHIVQAKIFLPLPESVDHDLGKISBITLAILGLPIHFVFP"

BASE COUNT 1084 a 676 c 634 g 923 t
ORIGIN

Query Match 32.0%; Score 900.2; DB 4; Length 3317;
Best Local Similarity 61.4%; Pred. No. 3,1e-239;
Matches 1635; Conservative 0; Mismatches 978; Indels 48; Gaps 10;

QY 5 TCACAGGAGATGTACAGCAATGGGGCAATTAAGATTCTGTGTCACTTGATTTCTC 64
DB 43 TTACGTAAACATGTCGCAAAATGTCGTCTGTAATGTTATCTGTCTCACTTGC 102
QY 65 ACCTCTGAAGGGGCGCTGAGTAATTCACATTCACCTGACCAACAATGGCTATGAG 124
DB 103 ATCTCTGCTCGG--AAATGAAAGTTCAATGTAATTAATTTGATTAATGGTATG 159
QY 125 GCATTGCTGTCATGCAATGCCCAATGTGCCAGAGATGAACAACCTGATCAACAATAA 184
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 130 GCATTGCTGTCATGCAATTAACCCCACTGTGCCAGAGATGAATAAACTCATTTGAAACATA 219
QY 135 AGACATGATGGTACCAGCATCTCTGTATCTGTGGAAGCTACAGAAAGGATTTAT 244
DB 220 AGGAATGTGTAACCTAAGCTTCTACTTACTTCACTGCTTCACTGCAAGAAAGATTTAT 279
QY 245 TCAGAAATGTTGCCATTTTGTATTCCTGGAACATGGAAGCAAAAGGCTGATATGAGAC 304
DB 280 TCAGAAATGTGAGCATTTTAAATTCATGACCTGGAAATCAAAATCTGAGTACTCATAC 339
QY 305 CAAGCTTGAGACCTACACAAAATGCTGATGTTCTGCTGAGTCTACTCTCCAGTA 364
DB 310 CAAGAAATGATCATATGATACCAAGGATGTGCTAATTCCTATTAATAATATG 399
QY 355 ATGATGAACCTTACACTGAGCATGGCACTGTGGAGAGAAAGGTTAAAGATCCAC 424
DB 430 GAGATGATCCCTATCTACATTCATATGGAAGGTGTGGGAAAAGAAAATATATATAC 459
QY 455 TCACCTCGATTTCTATTCAGAGAAAAGTTAGCTGAATATGACACACAGTAGGCAT 484
DB 480 TTACTCCAAACTCTTGTGACTAATAATTCACATCTATGGGTCGGAGGAGATAT 519
QY 485 TTGTGCATGAGTGGGCTCATCTAGCATGGGAGATTTTGCAGATCAATATATGATGCA 544
DB 530 TTGTCCATGAGTGGGCTCATCTAGCATGGGAGATTTTGCAGATCAATATATGATGCA 579
QY 545 AATCTACTATTC--AATGGAAGATACAGCATGATGATGTTACAGGATTTACTG 601
DB 580 CATCTATATTTCCAGAAAGAACTATTTGAAGCAAGATGTTCACTCATATATCTG 639
QY 612 GTACAAATGTAG--TAAAGAGTGTCAAGGAGCAGCTGTACCAACCAAGATGCAT 658
DB 640 GTATTAATGTGTTTCAAGAAATCCCTGAGGAGCAGTATTAACAATGATATGACAGC 699
QY 659 TCAATTAAGTAAACAGCATCTATGAAAAGATGATGATTTGTTCTCCAAATCCGCCGCA 718
DB 700 GTGACTCAGACAGGCTGTATGAAAGCAAAATGTACATTCCTCCAAAATATCCGCA 759
QY 719 CGGAGAGCTCTTAATATGTTTGCACAAATGTGATCTATAGTTGATATCTGTACAG 778
DB 780 CTGCAAGAGATTCATTTATGTTTATGCAAGCTTCATCTGTAGCTAATTTTGTACAG 819
QY 719 AACAAACCAACAGAGAGCTCCAAACAAAGCAAAATCAAAATGCAATCTCCGACGA 838
DB 830 AAAAAACCAACAAATCAAACTCCAAACCAAAACAAATGCAATGCAAGCAAAACGA 879
QY 889 CATGGAGTATCCGTATCTGTAGGACTTAAAGAAACCAATCCCTATGCA-----A 892
DB 830 CATGGAGTATCCGTATCTGTAGGACTTAAAGAAACCAATCCCTATGCAAGCAAAATGA 939
QY 813 CACAGCCCAAAATCCCACTTCATTCATTCGTCAGATTTGACAAAGAAATGTGTGTAG 952
DB 940 ATCCACGACATCATCTATTCATTTCTGTCAGAGTCCAAACAGGGGTATCTGTGG 999
QY 953 TCTTTGACAAATTCGGAAGCATGGGACTGTAAACCGCTCAATGACGTATCAACAG 1012
DB 1030 TACTTGATTAATCTGGAAGCATGCTGTGCAAGAACCGCTCTTCAATGATGAAGAG 1059
QY 1013 GCCAGCTTTCCTGTCGAGACATGTGAGCTGGGCTCGGTTGGGATGGTGCATTTG 1072
DB 1030 CAGAACTATCTATTCATTCATTAATTTGAAAGGATCTTTAGTGGGATGGTGCATTTG 1119
QY 1073 ACAGTGTGCCCATGTACAAAGTAACTCATATGATTAACAGTGGGAGTACAGGAGCA 1132
DB 1130 ACAGTGTGTGGAATTCAAATCATTTAACAAGATTAAGATGATATGTTTACCAA 1179
QY 1133 CACTGCTCAAAAGATTTACTTCAGACAGCTTCAGAGAGGAGCTCATCTGACGGGCTTC 1192
DB 1180 AGATTCACCGCAAAATGCTCTACAGTAACTATGCTGAACTTCAATTTGATGAGGCTCA 1239
QY 1113 GATCGCATTTACTGTGATTTGGAAGA--AATTCACAGTATGATGATGATTTGCTG 1249
DB 1240 AAGCAGATTCACGCAATTAATTCACAGTGTGACAGAGTACTTCTGTTCTGAAATCATAC 1299
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

OY 1250 TCGTACGGAGTGGGAAGACACATATTAAGTGGTCTTTAAACGAGCTCAACAAAGT 1309
 DB 1300 TATTAACTGATGGGAAGATTAATGAATTAATCATCTTGAAGATGTAACGAAGT 1359
 OY 1310 GTCCATCATCCACAGAGTGGCTTTGGGGCCCTCTGAGCTCAAGAACTAGAGAGTGT 1369
 DB 1360 GTCCATCATCCACAGAGTGGCTTTGGGGCCCTCTGAGCTCAAGAACTAGAGAGTGT 1419
 OY 1370 CCAAAATGACAGAGAGTTCACACATATCTTCAGATCAGTTCACAAATAGGCTCA 1429
 DB 1420 CAATATGACAGAGAGATTCGTTTTTTTTCATTAAGCATTA-----ACTGCTTGA 1473
 OY 1430 TGAATGCTTTTGGGCTTTTTCATCAGAAATGAGCTGTCTCTCAGCGCTCCATCCAGC 1489
 DB 1474 CTATGCTTTTCACTAGATTTTCACTGAGAGTGAAGTGAACATCAGCTCAGCAGGCTATTCAGT 1533
 OY 1490 TTGAGAGTAAAGGATTAACCTCCAGAAACGAGAGTGAATGATGAGACATGATCTGG 1549
 DB 1534 TGAAGAGCAAGGCTTGAATAATACAGAAAGAAAGAAAGCAAGTGAAGTCTGTAG 1593
 OY 1550 ACAGACCGTGGGAAGAGACATTTTCTTATCAGCTGAGACAGCAGCGCTCCGCAAA 1609
 DB 1594 ACAGTACAGTGGAAATGACATTTCTTTGTGATGACATGACAAATCAAAACAGAA 1653
 OY 1610 TCTTCTCTGGGATCCAGTGAAGAGCA-----AGTGGCTTTGTAGTGAACA 1660
 DB 1654 TTGTCTCCAAATCCAAAGAAAGAAATATTAACCTCGATTCAGAAAGATAGT 1713
 OY 1661 AAAACACAAATAGGCTCACTCCAAATCCAGCATTTGTTAGTTGGCATTTGCAAT 1720
 DB 1714 TAAATTTGATCTGCTGCTGCTCAATATCTGTAATGAGAGACAGTACTTGGACT 1773
 OY 1721 ACAGTCT-----GCAAGCAAGCTCACAAACCTTGACCTGAGTACGCTCCGCTG 1771
 DB 1774 ACAGCTTTCTAAATATCATGACAGCTCTAAATGCTACAGTACAGTACCATCTGAG 1833
 OY 1772 CGTCAATGCTTACCTGCTCTCAATTAAGTGAATCTTCCAAACGAAACAGACACAGCA 1831
 DB 1834 CAAGAAGTCTTACTATACCCCGAGTAATTCACACAGCTCATGAGTCAACATACACAC 1893
 OY 1832 AATTCGCCAGCGCTGATGATTTTGAATTTGCGCAAGAGCTCCCAATTTCTCA 1891
 DB 1894 AATATCTTCAATGCAATGATTTTGAATTTGCAATGCAATGCAATGCAATGCAATGCA 1953
 OY 1892 GGGCAGTGTACAGCGCTGATGATTTGCAATGCAATGCAATGCAATGCAATGCAATGCA 1951
 DB 1954 GAATCATGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2013
 OY 1952 TGAATATGAGAGGCTGATGATTTGCAATGCAATGCAATGCAATGCAATGCAATGCA 2011
 DB 2014 GGGCAATGAGTGTGAGTGTGATGATTTGCAATGCAATGCAATGCAATGCAATGCAATGCA 2073
 OY 2012 CAATTTATGACAGATGATGATTTGCAATGCAATGCAATGCAATGCAATGCAATGCA 2071
 DB 2074 CAATTTATGAGATGATGATTTGCAATGCAATGCAATGCAATGCAATGCAATGCA 2133
 OY 2072 CAGCCAGAGAGAGTGTGATGATTTGCAATGCAATGCAATGCAATGCAATGCAATGCA 2131
 DB 2134 CGGCTAGGCTTAATTTAAGCAACACAGCAAAAGTTCTATTTGTTCCAGGCTACGTTG 2193
 OY 2132 AGAATGATGAATATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2191
 DB 2194 AAAAGGTAATTAATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2253
 OY 2192 ACAAGAGAGTGTGATGATTTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2251
 DB 2254 CTAAATGAGAGAGTGTGATGATTTGCAATGCAATGCAATGCAATGCAATGCAATGCA 2313
 OY 2252 CAATGCTCTCCA---TACCTGATCTCTCCAGCTGCGCAATCAGCAGCTCAAGCGG 2308
 DB 2314 CTCTCTCTGATATCAACCTCTCTGCTTCCACCCAGTAAATATCAAGATCTTGAGGCTA 2373

OY 2309 AATTCAGGGGCGAGTCTATTAATCTGACCTGGACAGCTCTCGGAGTATTAATGACC 2368
 DB 2374 AGTCAAAAGAG---ATTATATCACTTATGAGCAGCCCGCTGCAATGCTCTAGATA 2430
 OY 2359 ATGAAACGCTCAACATATATCTTTGCAATTAAGTACAAATATTTCTGATCTGAGACA 2428
 DB 2431 AAGAAAGGCAACACAGTCAATTAATGAATTAAGTAAGTATTCATGATGCTGCAAGAG 2490
 OY 2439 AGTTCATGATCTCTCAAGTGAATCTGCTGCTGATCCCAAGAGGCAAGCTG 2488
 DB 2481 ATTTGCAATGCGACCTTTGATGATCTTCAATCTAATACCTAAGAGAGCGGATCA 2550
 OY 2439 AAGAAAGTCTTTTGTAAACAGAAACATTAATTTGAAATGCGCAGATCTTTCA 2548
 DB 2551 AAGAAATTTTGAATTAACCCAGAAACATTTTGAAGTAAAGTAAAGTAAAGTAAAGTAA 2610
 OY 2539 TTGCTATCAGGCTGTTGATTAAGTCAATGCAATGCAATGCAATGCAATGCAATGCA 2608
 DB 2611 TTTCAGTCAAGGCTATACAGAGCCCAATCTCAGAGGTTTCTACATATTTGACAG 2670
 OY 2605 TATCTTTTATTTCTCCAC 2629
 DB 2671 CAATCAATTTATTTCTCTAC 2691

RESULT 12
 AFI15852
 LOCUS
 DEFINITION
 Mus musculus endothelial chloride channel (Clca2) mRNA, complete cds.
 AFI15852
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 2765)
 Eblie, R.C. and Pauli, B.
 A murine endothelial homolog of Lu-ECAM-1 associated with lung melanoma metastasis
 Unpublished
 2 (bases 1 to 2765)
 Eblie, R.C.
 Direct Submission
 Submitted (21-DEC-1998) Molecular Medicine, Cornell University, VMC C4-146, Ithaca, NY 14853, USA
 Location/Qualifiers
 1..2765
 /organism="Mus musculus"
 /strain="C57/BL6"
 /db_xref="taxon:10090"
 1..2765
 /gene="Clca2"
 17..2725
 /gene="Clca2"
 /note="putative metastatic addressin"
 /product="endothelial chloride channel"
 /protein_id="AA047626.1"
 /db_xref="GI:12043705"
 /translation="MYPGLQVLLFLHLHLLQNTSSMHLNSNGEYVIAINPSVE
 DERLPSIKETWQSTLYFEASOGVRYRNTSLVPMWRKSEYILMRKRSYOKAD
 VIVADPHOHGDPTLYOGCGGGRGYTHHPNPLLTNNLRIGRGVPAVHNAHL
 RRGVDEYVNDPRYISKRNTTEATRCSASTIGKKYVHECGRSCVTRACRDRSKRL
 YEPKCTFIDKIDTAGASTMEQNLSVVEFCTENNHNAAEPNLQKMCNRSYDMVI
 KSAIDFQNSPPMRKGTAPPPTFSLKSRRYCVLDKSGSKEDRLIRNQDAEL
 YLQIYEKSNVGLVTFDSAAHIQNTLIRITSSSDQKITANLPQATGSGTICHIQ
 AGFOATISDOSTSGSEIVLIDGDENGISCEPAVSRGAILIHTALGPSARELET
 LSDMGGLREYVANKHVSLLDAFPRISSTSGSVOALQLESKAFVRGAWINSTPV
 VSYVGNDFPFIYTWYQKPEIILDPPKRTITSDPQDELINISARLIQITGTEG
 TWYSTIGTKSGLITWYVTRKSPMEELTAHNSQSTAYPSPSMIVYVSVGFL
 PVIGANVTALIEAEHGHVTLLEMDNGAGADLVAKNDGIYTRFTDYHGNGRYSKLVRY

OARKKRLSLRÖKNKSLYIPGYENKIVLNPDPDQOEAEIATEVEDFNKRVISGGS
 FTVSGAPDGDHARVPPSKYVLEAEFIDYIQLTWTAPGVLEDKGRARHRIIVSQ
 HPLGDEDFNNATLVNASSLIPKEAGSEKFEKFEKETKINDOLYLALQAYNAGL
 TSEVNIQAVFTSTLEDSISALGDNISIMSTVMGLAVIFNSILN"
 BASE COUNT 832 a 656 c 635 g 642 t
 ORIGIN

Query Match 30.8%; Score 865.4; DB 10; Length 2765;
 Best Local Similarity 60.9%; Pred. No. 1.5e-229;
 Matches 1580; Conservative 0; Mismatches 976; Indels 39; Gaps 9;

QY 91 TCACATATTCAGCTGACACAACTATGCTATGAAAGCATTTGCTGCAATGACCCCAAT 150
 DB TCCATGCTCATCTCAACAGCATGATGAGGGTGTGTGATGCCATTAACCCCACT 139
 QY 151 GTGCCAGAGATGAACACTCTATTCACAAATAAAGAGATGGTGCCAGCATCTCG 210
 DB GTGCCAGAGAGAGAAAGGCTCATCCAGCATTAAGAAATGTAATCAAGCTTCTACC 199
 QY 211 TATCTGTTTGAAGCTACAGAAAGCATTTTATTCAAAATGTTGCCATTTTGAATCT 270
 DB TACCTGTTTGAAGCCAGCAAGAGAGATTTTTCAGAGACATTAAGCATTTAGTCCCG 259
 QY 271 GAAACATGCAACAAAGGCTGACTATGTGACACAAACTTGAGACTACAAAATGCT 330
 DB ATGACCTGGAATCAAAATCTGACTTAATGACCAAAACGAAATCTGTACACAAAGCA 319
 QY 331 GATGTTCTGTTGCTGAGCTACTCTCTCAGGTAAATGATGAACCCATACAGAGATG 390
 DB GACGCTATGTTGCTGAGATCTCCTCAGGTAAATGATGAACCCATACAGAGATG 379
 QY 391 GGCACCTGTGAGAGAAAGGCTGAAGATCACTCCTGATTTCTGATTTGACAGAA 450
 DB GGCAGCTGTGAGAGAGAGAGAGATCACTCCTGATTTCTGATTTGACAGATG 439
 QY 451 AAGTTAGCTGATATGACACAAAGGCTATTTGCTGATGAGTGGCTATATCAGA 510
 DB AACTGCTGATATGAGACCCGAGCAGAGTCTTGTCCATGAGTGGCTATATCAGG 499
 QY 511 TGGGAGATATTTGACAGAGTCAATATGATGAGAAATTTCTACTTATGTC--CAATGGA 567
 DB TGGGAGATATTTGATGAGATATGATGAGTGAACGCGGCTTTCTACATTTCTTGAAGAA 559
 QY 568 ATACAGCAGATGATGATGTTGACAGAGTATTTACTGTACAAATGTAGTAAAGAGTGTAG 627
 DB ATAGAGAGAAAGGCTGCTCCGAGCATACAGGCAAGAAAGTGTGTCCAGAGTGTAG 619
 QY 628 GGAGGACCTGTATACCAAAAGATGACATTTCAATAAGTAAAGAGCTATGAAAA 687
 DB AGAGGACCTGTATGACAAAGGCTGCTCCGAGCATGAGAGCAGGCTGTATGAACCC 679
 QY 688 GGAATGATGATTTGTTCTCCCAATCCCGCAGAGAGAGGCTTCTATATGTTTGCACAA 747
 DB AAATGTACATTTATCCAGAGCAAAATACAGAGCTGGGGCTCATATATGTCTACGAA 739
 QY 748 CATGTTGATTTATGTTGATTTCTGTACAGAAACAAACACACAAAGAAAGCTTCAAC 807
 DB AACCTCAATTTCTGTGTTGAATTTTGCACAGAAATTAACCAATGACAGAGCCCAAC 799
 QY 808 AAGCAAAATCAAAATGCAATCTCCGAGAGCATGGGAAGTATCCGATTTCTAGAGAC 867
 DB CTACAAAACAAATGCAATGCAAGAGAGCATGTGGAGTATTAACGCGCTGCTCTAC 859
 QY 868 TTTAAGAAACCACTCTCTATG-----ACAACACAGCCACCAAAATCCACGCTCTCATG 921
 DB TTTGAGATTTCCCTCCATGAGAGAAACAGAGGCTCTCTCCACCTACATTTTTCATG 919
 QY 922 CTGCAATTTGACAAAGATTTGTGTTTGTAGTCTTGAACAATCTGGAAGATGGCGACT 981
 DB CTCAAGTCTCAAGAGGCTGAGTGTGTGTTGTGATTAATCTGGAAGATGAGCAAA 979
 QY 982 GGTAAACCCCTCAATGCACTGATACAGAGCCAGCTTTTCTGCTGACAGAGCTTGA 1041

QY 1042 CTGGGGTCTGGGTTGGATGCTGACATTTTGCAGCTGCTGCCCATGTACAAAGTACTC 1101
 DB TGGAGGAGCTCATCTGACAGCGGGCTGTGATGGGATTCATGATTA--GGAG 1218
 QY 1162 TCAGAGGAGCTCATCTGACAGCGGGCTGTGATGGGATTCATGATTA--GGAG 1218
 DB ACTGTGGAATCTCAATTTGCTGAGTCCAGGAGATTTCAAGCAATTAACCTCCAGT 1219
 QY 1219 AATATTCACATGATGATTCGAATTTGCTGCTGACAGGATGGGAGACACACTATA 1278
 DB GACAGAGCTCTCCGTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1279
 QY 1279 AGTGGTCTTTAAGAGAGTCAACAAAGTGGTGCATTCATCCACAGTGGCTTGGGG 1338
 DB AGTTCCTGCTTTGAGCGGCTCTCTGACAGCGGTGCTCATTCATCCACATGCTTGGGG 1339
 QY 1339 CCTCTGACGCTCAAGAACTAGAGAGGCTGTCCAAATAGACAGAGGTTTACAGACATAT 1398
 DB CCTTCTGCTGCCGAGAACTGGAGACTGTGCACAGCATGACAGAGAGGCTTCTGAT 1399
 QY 1399 GCTTACATCAAGTTCAACAAATGCTCATGATGCTTTTGGGCGCTTTCATACAGA 1458
 DB GCAACCAACATGTT-----AAGCAGCTTATCGAGGCTTACAGTATTCATCTTACA 1453
 QY 1453 AATGAGCTGTCTCTACAGCTCTCCATCCAGCTGTGAGATTAAGGATTAACCTCCAGAC 1518
 DB AGTGCACAGCTCTCCACAGAGCTCTGCAAGTTGAGAGCAAAAGCTTAAATGTACAGCA 1513
 QY 1519 AGCAGTGAATGATGACAGAGTATGTTGACAGACAGCGTGGGAAAGACATTTGTT 1578
 DB GGGGATGATTAACAGTACAGTACTGTGAGACAGTACAGTACAGTACAGTACAGT 1573
 QY 1579 CTATACCTGACAGAGAGGAGGCTCCCAATCTCTCTGAGATCCAGTGGACAGAG 1638
 DB GTTATACCTGACAGAGAGGAGGCTCCCAATCTCTCTGAGATCCAGTGGACAGAG 1633
 QY 1634 CA-----AGTGGCTTTGTAGTGGACAAACAAACAAATAGGCTTACCTCAATC 1689
 DB TATATACCTGAGATTTCCAGATGATGATCACTAAACATCCGCTGCTGCTCAATTA 1683
 QY 1690 CAAGGATTTGTAAGTTGACATTTGGAATACAGTCT--GCAAGCAAGCTCAACAA 1746
 DB CCGGACCTGACAGAGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1753
 QY 1767 TTAGCCCTGACGAGTCCGAGGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1806
 DB ATTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1813
 QY 1817 TCAAAACGAAAG 1866
 DB GCTACAGAGATGAG 1873
 QY 1867 GCGCAAGAGGCTCCCAATTTCTAGGGCAGTGTACAGAGGCTGATGATGATGATGAT 1926
 DB AGCCAGAGGCTTTTGGCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1933
 QY 1937 GGAAGAAAGATTTACCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 1986
 DB GAGATCAAGTCAAGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1993
 QY 1994 GACGCTCTACTCAAGGATTTTACAGACTATGACAGAGATGATGATGATGATGATGAT 2046
 DB GATGCAATCTACAGAGATTTTACAGATTTATGATGATGATGATGATGATGATGAT 2053
 QY 2057 GTGGGGCTGTGGAGAGATTAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2106

CLONE DISTRIBUTION: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAK Plate: 10 Row: 5.

FEATURES

SCIENCE

CD5

971 a 718 c 683 g 765 t

0

1. 3137

/organism="Mus musculus"

/db_xref="LocusID:80797"

/db_xref="taxon:10090"

/clone="MGC:5738 IMAGE:3583804"

/tissue="type=Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."

/clone_id="NCI_CGAP_Mam1"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

20. 2728

/codon_start=1

/product="chloride channel activated 2"

/protein_id="AAH08147.1"

/db_xref="GI:14198178"

/translation="MPEGLGVLEFLTLHLONTSSMYHLNSGVEGVVILNINSPPE
DERLPISEKEMVQASVLYFEASGGRVYFNRIISLVPMTWKSEEEYIMPKRESDYKAL
VIVADPHLOHDDPYTLQYGCGRGOYIHPTNPEFLTDNRILRTIGPRGRVYEHRAH
RMGVEDYNDRPYISRKNTLEATRCASITGKRVVHECGSGCVTRACRDSKTRT
YEPGCTIPKQIPAGASIMFMQNLNVEECTENNNNAEPLNQNKNRSTWDYV
KASADFONSPMRCTEAPRPPTFELLSRRVVCVLDGSSMSKDEKRLIMNDAEEL
VLTQIVKESVSGVLTVDLSAHLQNTYLIKITSSSDYCKITPLNQLQAGTSGICHGLD
AGPALTSSDOSTGSEFVLTLDEEDNGISCCPEAVSSGAIHTITLALGSAARELETL
LSDMTGRLREYANKHVSLLIDAFERISSTSGSVQALQLESKAPNRAQWINSYV
VDSVGVNDTFVITVIVQKEPITLIDDKRKYITISDFODEDLNRSALQLPGTAETG
TWYSITGTQSOLTTMTVTTTARSPTEPLATNAHMQSTAYPSRMIVNRYVSGFEL
PVLCANTALIEAHGHOVTLLEWDNGAGDVTYNDGIIYTFEYDHGNGRYSLKVRV
QARKNRKRLSLRQNKSLIYIGYENCKIYLNPRPDVQBEAIEATVEDFNRTVSGGS
FTVSGADPDGQHAHVPPSPVTLDEAFETIGDYIOLTMTPGKVLVDKGAHRYIIRVSG
HPLGLODDENNAITVWASSLIPKAGSEKTFPKETPMTKIANPOLVITAIQVNEAGI
TSEVSNIAQAVKFTSLEDSISALCADISALISMTYWGALVAFINSTLN"

[illegible]

QY 331 GATGTTGTTGGTGGAGTCTACTCCAGGATATGTAACCTACTACGTGAGCATG 390
 DP 333 GACGTCAGATTGGCGATCCCTACCTGCAACATGAGACAGCCCTACACCTTTAGTAT 382
 QY 331 GGCACCTGTGGAGAGAAAGGTGCCACCTCACTCGATTCAATTCACAGAAAA 450
 DP 333 GGACAGTGTGGGAGAGAGAGACGTACTATCACTTCCAACTTCTACTCACTGAT 442
 QY 431 AAGTTAGCTGAATATGAGCACCAAGGTAGGGCATTTGTCCATGAGTGGGCTCATTAAGA 510

D	443	AACCTGCGATCTATGACACCCGAGGACGAGACTCTTTGTCCATGAGTGGGCCATCTCCGG	502
O	511	TGGGGAGATTTTGACAGCTACAAATATGATGAGAAATTTCTACTTATC---CAATGGAAGA	567
D	503	TGGGAGAGTATTTGAGAGATTAACCTGGACCGGCGCTTCTACATTTCTACAAAGAACACT	562
O	568	ATACAGACAGTAAATGTTTCAGCAGGTATTTACTGTCACAAATGATGATTAAGAAAGTGCAG	627
D	563	ATAGAAAGACAGACAGTGCCTCCGCCAGCATTCACAGGCAGAAAGGTGGTCCACGAGTGTACG	622
O	628	GGAGCAGCTGTTACACCAAAAGATGCACATTCATTAAGTAAACAGGACTCTATGAAAAA	687
D	623	AAAGCAGAGCTGTGTACACAGGCGCGTCCGGCGGTGACTCCGAAGACACGCGTGTATGAACCC	682
O	688	GGATGTGATTTGTTCTCCATATCCCGCAGACGGGAGAAAGGCTCTATAATGTTTGCACAA	747
D	683	AAATGTACTATTATCCACAGCAAAATACAGCACTGGGGCTTCATTAATGTTTATGACAA	742
O	748	CATGTTGATTTATAGTTGAATTTCTGTACAGACAAACCAACAAAGAAAGCTCCCAAC	807
D	743	AACCTCATTTCTGTGTTGAATTTTGACAGAAATTAACCAATGACAGAAAGCCCAAC	802
O	808	AAGCAAAATCCAAAATGCAATCTCCGAAGCATGGAAGTGAATCCGTATTTCTAGAGAC	867
D	803	CTACAAAACAAAATGCAATGCAAGAAAGCAGGGGATGTAAATCAAGAGGCTGTGTGAC	862
O	868	TTTAAAGAAACCACTCTATG-----ACAACACAGCAGCACCAAAATCCACCTTCTCATTG	921
D	863	TTTCAAGAAATCCCTCCCATGAGAGAAACAGAACGCGCTCTCTCACACTCATTTTCATG	922
O	922	CTGCGAGATTTGGACAAAGATTTGTGTGTTAGTCTCTTACAAATCTGAGACATGGCAGCT	981
D	923	CTCAAGTCCAGAAAGCGAGT	982
O	982	GCTAACCCCTCAATCGACTGAATCAAGACAGGCGAGGTTTCTGTGCGACAGATGTGAG	104
D	983	GAAAGCCGCTTATTTGCGATGAATCAGACAGCAACACTGTACTTACTCAATTTGTGGAA	104
O	1042	CTGGGGTCTCTGGGTTGGGATGCTACATTTGATGACAGTCTGCCCATGTACAAATGTGACTC	110
D	1043	AAGGAGTCTATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	110
O	1102	ATACGATTAACAGTGGAGAGACAGGAGACACTGCGCAAAAAGATTACTCGACAGCT	116
D	1103	ATAAAAATATACAGACTAGTATGATGACTACCAAAAGATCAACGCAACCTCCCAACAGGCT	116
O	1162	TCAGGAGGGAGCTCCATCTGACGGGGCTTCGATCGGACTTTACTGTGATTA---GGAAG	121
D	1163	ACTGCTGGAACCTCAATTTGCGATGCACTCCAGCAGATTTTACAGGAAATTAACCTCCAGT	122
O	1219	AAATATCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	127
D	1223	GACACAGACCTTCGCGTTCTGAGATGCTATTTGTCACATATGGGAGAAATATGAGAAAT	128
O	1279	AGTGGGCTCTTAAACGAGGTCAAAACAAAGTGGTCCATTCATCCACACAGTCCGTTGGGG	133
D	1283	AGTCTCTCTTTGAAGCGCGTCTCGCAGGCGGTGCATATTCACACACATCGCTCTGGGG	134
O	1339	CCCTCTGCACTCAAGAACTAGAGAGACTGTCCAAAATGACAGAGGTTTACACAGATAT	139
D	1343	CTTCTTGTGTGCCAGAACTGGAACACTGTGCAGACATGTGACAGAGGCTTCTCTTCTAT	140
O	1399	GCTTCAGATCAAGTTTCAGAAACAAATGAGGCTCATTTGATGATGATGATGATGATGATGATG	145
D	1403	GCGAACAACAATG-----AAGACGCTTATTCAGCGCTTTTCAGTATGATTTCTATACAA	145
O	1459	AATGAGCTGTCTCTCAGCGCTCCATCCACACTTGAAGTAAAGGATTTAACCTCCAGAAC	151
D	1457	AGTGGCAGCTCTCCACAGACAGCTCTCAGATGTGAAGAGCAAAAGCTTTCATATGCAGAGA	151
O	1519	AGCCAGTGGATGATGGCAGACAGTATGCTGTGACACACCGTGGGAAAGACATTTGTTT	157
D	1517	GGGGATGATGATTAACAGATACGATACCTGTGTGAAGATACCGTGGCAGACAGCATGTTCTT	157

QY	1578	CTATTCACCTGGAGACAGCAGCGCTCCCAATCCCTCTCTGGGATTCACAGTGGACAGAG	1638
DB	1577		1637
QY	1576	GTTATCACTGGACGGTACAAAAGCCAGAAATATCTTTTCCAGATCCAAAAGAAAAA	1636
QY	1639	CA-----AGCTGGCTTTGTAGTGCACAAAACCAACCAAAATGGCTACCTCCAAATC	1689
DB	1638	TATATACCTCAGATTTTCCAAAGATGATGAACCTAAATCCGGTCTGTGCATTCAAATA	1686
QY	1636	CCAGCAATTCGTAAAGTTGGCACTTGGAAATACAGTCT---GCAGCAAGCTCACAAAC	1746
DB	1635		1745
QY	1634	CCGGGCATTCGACAGACAGGACTTGGACTTACAGATCACGGGTACCAAGTCTCAGTTG	1756
QY	1747	TTGACCCGACTGTACAGTCCGCCGCCGTCCAATGCAATCCCTGCTCCAAATACAGTGCT	1806
DB	1746	ATTACAAATGACAGTACACCTCGAGCAAGAAGTCCACCATGAAACCACTCCGTGTAC	1816
QY	1807	TTCCAAAAGCAACAGACACACAGCAAAATTTCCCGCCCTCTGTGTATGTATGCAAAATTT	1866
DB	1806	GCTCATGATGATCGAGCACACAGCCAGTACCTTACGCGGATATTTGTATGTACGGGTC	1876
QY	1853	CGCCAGAGACCTTCCCAATTTCTCAGGGCCAGTGTACAGCCCTCATTTGAATCAGTGAT	1926
DB	1852	AGCCAAAGGTTTTTGGCTGTCTTAGAGGCCAAATGTCCAGCCATATATGAAGCTAACAT	1936
QY	1947	GGAAAAACAGTTACTCTTGGAACTACAGTAATAGGACAGGTGTGATGATCTACTAAGAT	1986
DB	1937	GGACATCAAGTACCTCTGGACGACTGTGGACAAATGGGCGAGTGTGATCTGTTAAAT	1996
QY	1947	GACGGTGTACTCACTCAAGATATTTCCACAACTTATGACACGAATGTATGATACAGTATAA	2046
DB	1947	GATGGCACTCTCACAAATATCTTACAGATTTATCATGAAATGATATGATACAGCCTAAA	2056
QY	2007	GTCGGGGCTCTGGAGAGATTAACGACGCCAGACGAGAGTATACCCAGACAGTGGCA	2106
DB	2037	GTCGGTGTCCAGGCAAAACAAAAACAAAGCCAGACTAAGCTTAAAG---CAGAGAACAG	2113
QY	2107	GCACGTGCATACCTGTGCTGATTTGAAATGATGAATATCAATGTGATATCCACCAAGCT	2166
DB	2104	TCTTTATATATACCTGGCTATGTGGAAATGTGTAATTTGTACTCAATTCACCCAGACCA	2173
QY	2107	GAAATTAATTAAGATGATGTTCACACACAGCAAGTGTGTTTCAGCAGAACTCTCGGGA	2226
DB	2104	GATGTCCACAGAAAGGCAATGAAAGCTACAGTGAAGACTTCAACAGATTAACCTGGGA	2233
QY	2227	GGCTCATTTGTGGCTTCTGTATGTGTCCCAATTCCTCCATACCTGATC---TCTTCCCACT	2283
DB	2204	GGGTCAATTTACTGTGTGTGAGGGCCCTCGATGGGACCAAGCTCGTGTCTCCACCA	2293
QY	2204	GGCAAAATACCGACCTGAGAGGGGGAATTCACGGGGGAGTCTATTATCTGACTGG	2343
DB	2204	AGTAAAGTCACAGAGCTGGAGGGCTGAGTTTATAGG---ATTATATTCAACTTACATGG	2350
QY	2304	ACAGCTCTCGGGGATGATTAATGACCAATGGAACAGCTCACAGATATATTCGATTAAGT	2403
DB	2301	ACGGCCCTCTGGCAAGGTTCTCGATTAAGGAAGGCGCATATGATCATCAGAGTGGAC	2410
QY	2404	ACAAGTATCTGTATCTTCAGAGACAGTTCATGAATCTCTTCAAGTCAATACACTGCT	2463
DB	2401	CAGCATCTCTGGGTCTCCAGAAAGTTTAAACATGTCACTTATGTAAGTCTTCAGT	2470
QY	2404	CTCATCCCAAGGAGGCAACTCTGAGAGAGTCTTTTGTATTAACCGAGAAACATTTCT	2523
DB	2401	CTGATACTTAAGAGGCTGGCTTAAAAGAAATTTAATTAATTCAAACCGAAACCTTTTAA	2530
QY	2504	TTTGAAAATGGACAGATCTTTTCATTCGTATTCAGGGCTGTGATATAGCTGATCTAAA	2583
DB	2501	ATACCAATATGACACCAAGCTCTCATTTGCAATTCAGGACATCAATATGAAGTGTCTCAC	2590
QY	2504	TGCAAAATATCCAACTTGCACAGATATCTTTGTTTATTTCTCCACAGACTCCGCCAG	2643
DB	2501	TCTAGGCTTCCAAACTTGACACAGGCTGTCAAGTTCATTTACTTCTTGAAGATATGATCTCT	2650

QY	2644	ACACCTAGTCTCTGAT	2658
Db	2651	GCACTGGGTGCTGAT	2665

RESULT	14
AF108501	
LOCUS	
DEFINITION	AF108501 3058 bp mRNA
ACCESSION	Mus musculus Ca(2+)-sensitive chloride channel 2 (cacc) mRNA.
VERSION	AF108501
KEYWORDS	AF108501.1 GI:6465944
SOURCE	house mouse.
ORGANISM	Mus musculus

BASE COUNT	951 a .	706 c	670 g	731 t
ORIGIN				

Query Match	30.7%;	Score 863.8;	DB 10;	Length 3058;
Best Local Similarity	60.8%;	Pred. No. 4.2e-229;		
Matches 1579;	Conservative	0;	Mismatches 977;	Indels 39; Gaps 9

```
QY      TCACCTATTGAGCTGACACMACATGGGTATGAAGGCATTGTCGTTGCAAATCGACCCCAAT    150
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      TCCAATGGTGCATCTCAACAGCATATGATATACGAGGCTGTGGTCATTGGCCATTAAACCCCAAGT    139
```

111 GTGGCGAAGATGTAACACATCATTTCAACAAATTAAGACATGTGTACCCAGGACATCTGTG 210
112 GTGCGAAGGACGAAAGGCTCATTCGCCAAGCATTAAGGAATGGTAATCACTCAAGCTTCTAC 199
113 GTGCGAAGGACGAAAGGCTCATTCGCCAAGCATTAAGGAATGGTAATCACTCAAGCTTCTAC 199
114 TATCGTTTGAAGCTACAGGAAGGCAATTTTATTTCAAAAATGTTGCCATTTTGATTTCT 270
115 TACCTGTTTGAAGCCACCACCAAGGAGAGTTTATTTGAGAACATTAAGCATTTATGTCGG 259
116 GAAACATGGAAGACAAAGGCTGACTATGTGAGACCACAACTTGAGACCTACACAAAATCT 330
117 ATGACCTGGGAAGTCAAAAATGTGATCTTAATGCCAAAACGAAATCTGACACACAACCA 319
118 GATGTCTGTTGCTGTGAGTCTACTCTCCAGTAATGATGAACCTTACACTGACGAGATG 390
119 GACGTCATAGTTGGGATTCCTCACTGACAAACATGAGAGAGACCCTTACACCTTCAGAT 379
120 GCGCACTGTGAGAGAAAGGTGAAGAGATCCACTCACTCTGTGATTTCAATTCAGGAAAA 450
121 GGACAGTGTGGGGACACAGGACAGTACATCACTTCACACAACTTCTTACTCACTGAT 439
122 AAGTTAGTGAATTTGGACCAACAAGTAGGGGCAATTTGTCCATGAGTGGGCTCATATAGA 510
123 AACTTGCTATCTATGACACCCCGAGGACAGTCTTTGTCCATGAGTGGGCCATCTCCGG 499
124 TGGGAGTATTTGACGAGTACAAATATGATGAGAAAAATTTCTATTTC ---CAATGAAGA 567
125 TGGGAGTATTTGATGAGTATTAACGTGACCGGCTTTCTACATTTCTGAAAGAACT 559
126 ATACAGAGATTAACATGTTCAAGCAGATTAATCTGTACAAATGTAGTAAGAAGCTGAG 627
127 ATAGAGACAAACAGTCTCCGACAGATCACAGCAACAAAGGTGTCCACAGTCTGAG 619
128 GGAGCAGCTGTTACACCAAAAGATGACATTTCAATTAAGTAAACAGACTATGAAAAA 687
129 AGAGGACACTGTGTGACAAAGGGGTGGCGGCTGATCGAAGCAAGCGCTGATGACACC 679
130 GGATGTGATTTGTTCTCAATCCCGCAGACGGAGGAGGCTTCTATTAATGTTTGCACA 747
131 AATGTACATTTATCCAGCAAAATTAACACAGCTGGGGCTTCATTAATGTTATGAA 739
132 CATGTGATTTATAGTTGAATTCGTACAGCAAAACACACAAAGAACCTGCTCAAC 807
133 AACTCATTTGTGTGTTGAATTTTGCACAGAAAATTAACACAAATGACAAAGCCCCAAC 799
134 AAGCAAAATCAAAAATGCAATCTCCGAGCACATGGAAGTAGTCGTGATTCGAGAC 867
135 CTACAAACAAATAGTCAATCGACAGACAGCTGGGATGTAAATCAAGCGCTCTGAC 859
136 TTTAAGAAAACCACTCCTATG ----ACACACACCCACCAATGCCACTTCTCATGTG 921
137 TTTCAAAATTCCTCCCATGAGAGAACGAAAGCGCTCTCCACATCAATTTTCATTTG 919
138 CTGCGATTTGACAAAAATTTGTGTGTTAGTCTTTGACAAATTCGAGAGATGCGACT 981
139 CTCAAGCTCCAAAGCGAGTGTGTGTGTTGCTGTGATTAATCTGGAAGCATGGAACAA 979
140 GGTAAACGCCCTCAATCGACATGATCAAGCAGCCAGCTTTTCTGTGTCAGACATGTAG 1041
141 GAAAGCCCTCTTATTCGAATGAATCAAGCAGCAAGCACTGTACTTAATCTAAATTTGTGAA 1039
142 CTGGGCTCTGGTGTGGATGATGATCAATTTGACAGTGTGCCCATGTAACAAAGTAACTC 1101
143 AAGGAGTCTATGTTGGATTAATGACATTTTGAACAGGCTGTCCACATCCMAAATATATCTA 1099
144 ATACAGATTAACAGTGCAGTACAGGAGCACATCCCAAAAGATTAATCTGACAGACT 1161
145 ATAAAAAATTAAGAGTACTAGTACATCAAAAGATCAACGCAAACTCCCCCAACAGCT 1159
146 TCAGGAGGAGTCACTCATCTGAGGGGCTTCGATGGCATTTACTGTGTA ---GGAAG 1218
147 ATGTGTGAACCTTAATTTGTCATGTGACCTCCAGCAGATTCAGCAATTAACCTTCAGT 1219

QY	1219	AAATATCACTAGTATGGATATGGTAATTTGGCTGTACAGATGGGGAAGCAACAACCTATA	1278
Db	1220	GACCAAGACACTTCCGGTTGTGAGATCTGATTCTCAGATGGGGAATATGCAATA	1279
QY	1279	AGTGGGCTTTTAAAGAGTCAAAAGTGTGCATCATCCACACAGTCGTTTGGG	1338
Db	1280	AGTTCTCGTTTGGCGCGTCTCCAGGGGTCATCATCAACCATCCGTCGTGGG	1339
QY	1339	CCCTCTGACAGTCAAGAACTAGAGAGCGTGTCCAAATGACAGAGGTTTACAGCATAT	1398
Db	1340	CTTTCTCTGCGGAGACTGGAGACTCTGTGACATATCAGAGAGGCTTGTTTCTAT	1399
QY	1399	GCTTTCAGATCAAGTTTGAGAACATTTGGCTTATGTATGCTTTTGGGGCCCTTTCATCAGGA	1458
Db	1400	GCCAAACAACATGT - - - - - AAGCACCTTATGACGCTTTCAGTAGATTTCACTATCA	1453
QY	1459	AATGAGAGCTGTCTCTACGCGCTCCATTCACAGTTGAGAGTAAAGGATTAACCTCCAGAC	1518
Db	1454	AGTGGACGCGTCTCCAGCGGCTGTGCAATTTGGAGACAAAGCCCTTCATATCTCAGACA	1513
QY	1519	AGCCAGAGGATGAATGACAGATGATCGTGAGACAGCCTGGGAAAGACACTTGT	1578
Db	1514	GGGGCATGATAAACAGTACAGTACTTTGTGTGACATGATCCGTGGCAGACAGCATTTCTT	1573
QY	1579	CTTATCACCTTGACACACGACGCTCCCAATTCCTCTGTGGAAATCCAGTGACAGAG	1638
Db	1574	GTTTATCACCTGGAGCGATGACAAAGCCAGAAATCTTCTCAAGATTCACAAAGAAAAA	1633
QY	1639	CA - - - - - AGGTGGCTTTGTAGTGAGCAAAACACCAAAATGGCCCTACCTCCCAATC	1683
Db	1634	TATTTAACTCAGATTTTCCAAAGTATGATCTTAACATCCGCTGTCTGTGACTTCAATA	1693
QY	1690	CCAGGACTTGTCTAAGTGTGGCACTTGGAAATACAGTCT - - - GCAAGACAGCTCACAAAC	1746
Db	1694	CCGGGCACTGGAGAGACAGGTACTTTGGACTTACAGCATCACGGTATCCAAAGTCTGATTG	1753
QY	1747	TTTGACCTGAGTGTACAGTCCGCTGGTCCAAATGCTAACCTCGCTCCATTTACAGTACT	1806
Db	1754	ATTTCAAATGACAGTACGACCTCTCAGCAAGAAAGTCCACCATGCAACCACTCTGCTATCC	1813
QY	1807	TCCAAAACGAAACAGACGACACAGCAAAATTTCCCGAGCCTCTGTGATTTATCAAAATTT	1866
Db	1814	GCTCAGATGAGTGTGAGACACAGCCAGTACGATCCGTAGCGGATGATTGTGATGTACGGGTG	1873
QY	1867	CGCCAAGAGAGCTCCCAATTTCTCAGGGCCAGTGTACAGCCCTGATTGATACAGTAAAT	1926
Db	1874	AGCCAAAGGTTTTCCTGTTTGTAGAGAGCGGATGTACAGCCATATMAAACCTCAACAT	1933
QY	1927	GGAAAAACAGTTACTTGGAACTACTGAGTAATGAGAGCGGTGTGATCTCTCAAGAT	1986
Db	1934	GGAATCAAGTACACCTGTGAGCTATGTGGACATATGGGGCAGGTGTCTATCTTTAAAT	1993
QY	1987	GAGGCTGTCTATCTCAAGGTATTTTACAACTTATGACAGCAATGTAGATACAGTATAAA	2048
Db	1994	GATGGCACTTGTACACAAATACTTTACAGATTAATCATGAAAAATGTGTAGATACAGCTAAA	2053
QY	2047	GTGGGGGCTCGGGAGAGATTAAAGCAGCCAGAGCGGAGAGTATACCCAGAGAGTGA	2106
Db	2054	GTGGGTGCCAGGCAAAAAAAGGAGGAGCTAAGCTTAAAG - - - CAGAGAGAACAG	2110
QY	2107	GCACTGTACATACCTGGCTGGATTGTGAAATGATGAATATCAATGTGAAATCCACAGAGCT	2166
Db	2111	TCTTTATATATACCTGTGCTATGTGTGAAAAATTTGTACTGAAATCCACCGAGACA	2170
QY	2167	GAAATTTAATGATGATGTTTCAACACACAGCAAGTGTGTTTACGACGAATCTCTGGGA	2228
Db	2171	GATGTCCAAAGAAAGCAATGAAAGCTATCAGTGTGAGAGACTTTCACAGAGATTAACCTCTGGA	2230
QY	2227	GGGCTCATTT - - - TGGGGCTTCGATGTGCCAAATAGCTCCCATTCAGTCGATCTCTGCCACT	2283
Db	2231	GGGTGATTTGAGTGTGTGTGGAGGGCCCCCTGATGTGGGACATATGTCTGTGTGTCCACACA	2293
QY	2284	GGCCAAATCACCGCACTGAAGGCGGAATTTACGGGGGCAATCTCATTTATCTGACTTGG	2343

D6	2231	AGTAAAGTCACAGACCTTGAGGCTGAGTTTAGTG---	ATTACATTCACATTAACAAGG	2347
OY	2334	ACAGCTCCTGGGGATGATTAATGACCATGAACAGCTCAACAGTATATCATTCGAATAACT	2403	
D6	2338	ACGGCCCTCGGCAGAGGTCTCGATTAAGGAAGAGCGCATAGATCATCATAGAGGTGAGC	2407	
OY	2434	ACAAATATCTTGATCTGAGAGACAAAGTTCATGAACTCATCTCAAGTCAATATCACTGCT	2463	
D6	2438	CAGCATCTCTGGGCTCTCCAAAGAATTTTAACAATGTTACTTTAGTGAATAGCTTCAGT	2467	
OY	2474	CTCATGCCAAGGAAGCACCACTCTGAGGAAGTCTTTTTGTTTAAACGAAAACATTAAT	2523	
D6	2478	CTGATACCTTAAGAGGCTGGCTCAAANAAGAACATTTTAATTCACAGAAACCTTTTAA	2527	
OY	2524	TTTGAAATGGCACAGATCTTTTCATTTCTCATTTTCAGGCTGTTGATTAAGTGCATCTGAA	2583	
D6	2538	ATVGAAGAAATGACACCCAGCTCTACATTTCAATCCAGGATACAAAGAGCTGCTTCACG	2587	
OY	2534	TCGAAAAATATCCAAATTCAGACGAATATCTTTTATTTCTCCACAGACTCCGCCAAG	2643	
D6	2538	TCTGAGGCTCTCAACATTTGCACAGGCTGTCAAGATTACTCTCTGAGAAATAGATATCT	2647	
OY	2634	ACACCTAGTCTGAT	2658	
D6	2638	GCACTGGGTGCTGAT	2662	
RESULT	15			
LOCUS	AF047838	3022 bp	mRNA	R0D 26-NOV-1998
DEFINITION	Mus musculus calcium-sensitive chloride conductance protein-1 (mClCa1) mRNA, complete cds.			
ACCESSION	AF047838			
VERSION	AF047838.1	GI:3925280		
KEYWORDS	house mouse.			
SOURCE	Mus musculus			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 3022)			
AUTHORS	Gandhi,R., Ebble,R.C., Gruber,A.D., Schreuer,K.D., Ji,H.L., Fuller,C.M. and Pauli,B.U.			
TITLE	Molecular and functional characterization of a calcium-sensitive chloride channel from mouse lung			
JOURNAL	J. Biol. Chem. 273 (48), 32096-32101 (1998)			
MEDIA	99041980			
REFERENCE	2 (bases 1 to 3022)			
AUTHORS	Gandhi,R., Ebble,R.C., Gruber,A.D. and Pauli,B.U.			
TITLE	Direct Submission			
JOURNAL	Submitted (12-FEB-1998) Vet. Pathology, Cornell Univ., VMC C4-146, Ithaca, NY 14856, USA			
FEATURES	Location/Qualifiers			
SCIENCE	1..3022			
gene	/organism="Mus musculus"			
CDS	/db_xref="taxon:10090"			
	/tissue_type="whole lung"			
	1..3022			
	/gene="mClCa1"			
	18..2726			
	/gene="mClCa1"			
	/note="calcium-activated chloride channel; Alternative chloride channel to CFTR that may influence severity of Cystic Fibrosis in different tissues; mCaCC-1"			
	/codon_start=1			
	/product="calcium-sensitive chloride conductance protein-1"			
	/protein_id="AAC79982.1"			
	/db_xref="gi:3925281"			
	/translation="MPVGIOVLFLTLHLONTSSWVHNSNGEAVIAINPSVE DERLISIKEMVAQSYLYLEPAQSGRYFENSTILYPMWKSKREYLMRESYDAAD VIYADPHLOHDDPYITLQGCGRGQYIHETPNFLTLDNRITGPGRGVFHEMAHL RMGEVDYNODPEYMSRKNTIEGRTSTRITGNVVHNCEKNCVPRACKRDSKTRL			

YEPKCTIPDKIQTAGASIMFMONINSVERCTEKNHNAEPNIONKNCNRSTWDVI
KTSADFOAPMRGTEAPRPTFSLKRRRVCLVLDRKSGMEDLIIRNMQAEIL
YLTOIVEKSEWGLVTFEDSAHIONYLKIRISSPYQKITALPOAAGSISICGL
QFOATISSDOSTSGEIVLLDGEDNIRSEFVSGSALIHITAGSARLEET
LSDMNGGIRFANDNLIDAFRISSTSSVSQALOBESKADYVAGNINTV
IDSTVNDTFVITMVKPEIILDDPRGKTTSDPDOKNINISAKIOLPETAETG
TWTYSITGKSQLITMIVITVTLNRSPTMPLRSDSTAOYPRMIVYARVSGFL
PVLGANVTALIEHGHQVTELEMDNGAGLTVKMDGJYTRFYDHNGRSLKVR
OARNRKLSLRKRSKSLYIPGYENGLNPPRDOEALIEATVEDFNSTGGS
FTVSGAPDGBHARVPEPSKYTDLEAEKIGYHILWTAPGKVLNGBAHRIIMSO
HPIDLOEDFNNAITLVNASLIPKENGSETEKRETKRINGIOLYIATIQADNASL
TSEVSNIAQVKTLSLEDSIALGDSDISAITMGLVITFENSIILN*

BASE COUNT 923 a 710 c 661 g 728 t
ORIGIN

Query Match 30.7% Score 863; DB 10; Length 3022;
Best Local Similarity 61.1%; Pred. No. 6,9e-229;
Matches 1568; Conservative 0; Mismatches 960; Indels 39; Gaps 9;

91 TCACATCATCTGACCTGACACAAACATGGCTATGAAGCATGTGTTGCATGACCCCAT 150
81 TCACATGTCATCTGACACGACATGATGAGGCTGTGTCATTCATTAACCCCAT 140
151 GTGCCAAGATGAAGACACTCATCAACAATPAAGACATGGTACCAGCATCTCTG 210
141 GTGCCGAGAGCAGAAAGCTCATCCCAAGCATAAAGAAATGATCTCAAGCCTTACC 200
211 TATCTGTTGAAGCTACAGGAAGCATTTTTCAAAATGTTGCCATTTGATCTCT 270
201 TACCTGTTGAAGCAGCCAGAGAGTTTATTCAGAAATAGACATTAATAGCCCG 260
271 GAACATGGAAGACAAAGCTGACTATGTGAGACCAAACTGAGACCTACAAAATGCT 330
261 ATGACCTGGAATCAAAACCTGATCTTAATGCCAAAACGAGATCTGATGACAAAGA 320
331 GATGTTCTGCTGCTGAGTCTACTCCTCCAGTAAATGATGAACCCCTACACTGAGCATG 390
321 GACGTCATGTTGGGATCTCACTCGCAACATGAGAGACCCCTACACCCCTGAT 380
391 GGCACTGTGGAGAGAAAGGTGAAGGATCCACCTCACCTCCGATTTCTACTGAGGAAA 450
381 GGACAGTGTGGGACAGAGACGATACATCACTTCACTCAAACTCTCTACTGAT 440
451 AAGTATGCTGATATGACACAGGATGAGGATTTGCTCATGAGTGGCTATCAGCA 510
441 AACTTGGCATCTACGAGACCCGAGGACAGCTTTGTCTCATGAGTGGCCCATCTCCG 500
511 TGGGAGATATTTGACGAGTACAAATAATGAGAAATCTTACTTATC---CAATGGAAGA 567
501 TGGGAGATATTTGATGAGATATACGAGGACGCTTCTACATGCTCGAAGAAACACT 560
568 ATACAAGCAGTAAAGTGTACAGAGTATTACTGTACAAAGTACTAAGAAAGTGTAG 627
561 ATACAAGCAGCAGGATGTTCACACAGGATCTGCGACCAATGTGTTACAACTGTGAG 620
628 GGAGGACAGCTTACACCAAAAGATGACATTCATTAAGTAACAGGACTATGAAAA 687
621 AGAGCAGTGTGTACCAAGGGCGTGGCGGTGACTGACAGACAGCGTGTATGACCC 680
688 GATGTGATGTTGTTCTCCATCCGCGACGAGAGAGGCTTCTAATAATGTTTGCACA 747
681 AATATACATTTATCCACAGCAAAATATACAGAGCGCTGGCGCTCTCAATATGTTACGACA 740
748 CATGTTGATTTCTATAGTGAATCTGTACAGAACAAACACCAACAAAGACCTCCAAAC 807
741 AACCTTAATTTCTGTGTTGATTTTGCAGAAAAAACCAACCAACCAAGCCCAAC 800
808 AAGCAAAATCAAAAATGCAATCTCCGAGCAGCATGGAGTATCGGTATCTGAGAC 867
801 CTACAAAACAAAATGTGATGCGAAGCAGTGGAGTATACAGAGCTCTGCTGAC 860
868 TTTTAAAGAAACCACTCTATG-----ACACACAGCCACCAAAATCCACCTTCTCATTTG 921

Db	1935	GGACATCTCAAGTCACCTTTGGAGCTGTGGGACAAATGGGGCAGGGTGTATCTGTAAAAAT	1994
Oy	1987	GACGGTGTCTACTCAAGATATTTCACACTTATGACACGAATGGTAGATACAGTAAAA	2048
Db	1995	GATGGCACTATCACAAACATCTTTACAGATTATCAAGAAAGTAGATACAGCTTAAAA	2054
Oy	2047	GTGGGGGCTGTGGGAGGATTTAAAGCAGCCAGACGGAGAGTAGTATCCAGACAGTGGGA	2106
Db	2055	GTGGCTGTCCAGGCGACAAAGAAACCAACAGACTAGCTTAAAG---CAGAAAGAACAG	2111
Oy	2107	GCACGTACATACCTGCTGGATTGAGATGATGAATCAATGGAATCCACCAAGACCT	2166
Db	2112	TCTTTATATATACCTGGCTATGTGGAAAAATGGTAAATATGTACTGAATCCACCCAGACCA	2171
Oy	2167	GAATTTATATAGATGTGATTTCAACACAAACCAAGTGTGTTACGACGAACATCCTCGGGA	2226
Db	2172	GATGTCACGAAGAGGACATATGAAAGCTACAGTGGAAAGACTTTAACAGATTAACCTTGGGA	2231
Oy	2227	GGCTCATTT---TGTGGCTTCTGATGTGCCAAATGCTCCATACCTGATCTCTCCACCT	2283
Db	2232	GGGTGCTTTACTGTGTCTGGAGGGCCCCCTGATGGGAGACAGCTCTGTGTCCACCA	2291
Oy	2284	GGCCAAATCACCGACTGGAAGGGGGAATTCACGGGGGCAGTCTATTAATCTGACTTGG	2343
Db	2292	AGTAAAGTCACAGCTCGGGGGCTGATTTAAGGTGAATTA---ATTCACTTTCATCAAGG	2348
Oy	2344	ACAGCTCTGGGGATGATATATGACCAATGGAACAGCTCACAAATATATCAATTCGATTAAGT	2403
Db	2349	ACGGCCCTGGCAGAGTGTCTCGAAGTGGAAAGGACACATAGATACATCATCGAATGAC	2408
Oy	2404	ACAAATATCTTGATCTCAGAGACAAAGTCAATGAATCTCTTCAAGTAGAATCTACTGCT	2463
Db	2409	CAGCATCTCTGGATCTCCCAAGAAATTTTAAACAAAGCTACTTGTAGTAGATCTCTCCAGT	2468
Oy	2464	CTCAATCCAAAGAGAACCACTCTGAGGAAGTCTTTTGTAAACCGAAAAACATTAAT	2523
Db	2469	CTGATACCTTAAGAGAGCTGGCTCAAAAAGAAACATTAAATTCAAAACGAAACTTTTAAA	2528
Oy	2524	TTTGAATAATGGCAGACAGATCTTTTCATTGCTATTCAGAGCTGTGATTAAGTGCATCGAATA	2583
Db	2529	ATACCAATATGGCAATCGAGCTCTCAATTCGAATCAACGACGACAGAAATGAAGCCAGTGCAC	2588
Oy	2584	TCGAATAATATCAAACTTGAACGAGATATCTTTGTTATATCTCCACAA	2630
Db	2589	TCCTAAGGCTTCCAACTATCGACAGGGCTGCACAACTTACTTCTCTGTGA	2635

Search completed: April 4, 2002, 03:11:15
Job time: 60957 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 09:17:22 ; Search time 521.98 Seconds

(Without alignments)
96.756 Million cell updates/sec

Title: US-09-049-696-1

Sequence: 1 GAAATCAGGAGGAGATGTAC.....ATCTCTGATCTGTTGAAG 223

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents, NA:*

1: /cgn2_6/pdata/1/lna/5A_COMB.seq:*
2: /cgn2_6/pdata/1/lna/5B_COMB.seq:*
3: /cgn2_6/pdata/1/lna/6A_COMB.seq:*
4: /cgn2_6/pdata/1/lna/6B_COMB.seq:*
5: /cgn2_6/pdata/1/lna/PCTUS_COMB.seq:*
6: /cgn2_6/pdata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	51.6	401	4	US-09-221-298-34
2	33	14.8	6924	1	US-08-015-973-2
3	33	14.8	6924	2	US-08-448-164-2
4	31.4	14.1	1876	2	US-08-466-589-7
5	31.4	14.1	1876	2	US-08-700-636-7
6	31.4	14.1	1876	3	US-08-467-574-7
7	31.4	14.1	1876	4	US-09-217-345-7
8	29	13.0	7032	4	US-09-324-867-1
9	27.8	12.5	1438	4	US-09-187-331-4
10	27.6	12.4	3663	4	US-09-499-884-11
11	27.6	12.4	12720	1	US-08-403-866-11
12	27.4	12.3	1620	6	5449756-10
13	27.4	12.3	2306	6	5198359-3
14	27.4	12.3	2306	6	5449756-3
15	27.2	12.2	6960	6	5449756-3
16	27.2	12.2	8176	2	US-08-841-349-3
17	26.8	12.0	549	2	US-08-851-190-2
18	26.8	12.0	2364	2	US-08-838-2198-5
19	26.8	12.0	2364	3	US-09-233-336A-5
20	26.8	12.0	2364	3	US-09-233-752A-5
21	26.8	12.0	2364	4	US-09-402-036-5
22	26.8	12.0	2612	1	US-08-471-033-31
23	26.8	12.0	2612	2	US-08-471-044-31
24	26.8	12.0	2612	2	US-08-463-483A-31
25	26.8	12.0	2612	2	US-08-471-046A-31
26	26.8	12.0	2612	2	US-08-470-566B-31
27	26.8	12.0	2612	2	US-08-838-219B-3

28	26.8	12.0	2612	2	US-08-469-334-31	Sequence 31, Appl
29	26.8	12.0	2612	3	US-09-300-529-31	Sequence 31, Appl
30	26.8	12.0	2612	3	US-09-233-336A-3	Sequence 3, Appl
31	26.8	12.0	2612	3	US-09-233-752A-3	Sequence 3, Appl
32	26.8	12.0	2612	4	US-09-402-036-3	Sequence 3, Appl
33	26.6	11.9	1299	1	US-07-688-352C-17	Sequence 17, Appl
34	26.6	11.9	1299	2	US-08-474-379C-17	Sequence 17, Appl
35	26.6	11.9	1299	3	US-09-146-249A-17	Sequence 17, Appl
36	26.6	11.9	1299	3	US-08-206-188B-17	Sequence 17, Appl
37	26.6	11.9	1299	5	PCT-US91-02714-11	Sequence 17, Appl
38	26.6	11.9	1299	5	PCT-US81-02714-17	Sequence 17, Appl
39	26.6	11.9	2370	4	US-08-960-780-5	Sequence 5, Appl
40	26.6	11.9	2370	4	US-09-073-898-5	Sequence 5, Appl
41	26.6	11.9	8584	1	US-08-358-160-66	Sequence 66, Appl
42	26.6	11.9	8590	1	US-08-358-160-70	Sequence 70, Appl
43	26.4	11.8	1838	3	US-08-848-564-1	Sequence 1, Appl
44	26.4	11.8	2022	1	US-08-803-973-6	Sequence 6, Appl
45	26.4	11.8	2022	1	US-08-803-972-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-221-298-34
Sequence 34, Application US/09221298
Patent No. 6284241
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.471
CURRENT APPLICATION NUMBER: US/09/221, 298
CURRENT FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 112
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 34
LENGTH: 401
TYPE: DNA
ORGANISM: Human
US-09-221-298-34

Query Match 51.6%; Score 115; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 5e-29;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-221-298-34
Sequence 34, Application US/09221298
Patent No. 6284241
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.471
CURRENT APPLICATION NUMBER: US/09/221, 298
CURRENT FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 112
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 34
LENGTH: 401
TYPE: DNA
ORGANISM: Human
US-09-221-298-34

US-08-015-973-2
Sequence 2, Application US/08015973
Patent No. 5604394
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHORYOSINE
WATER OF INVENTION: PHOSPHATASE-BETA
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:


```

      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patent Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/015, 973
        FILING DATE: 10-FEB-1993
        CLASSIFICATION: 435
        ATTORNEY/AGENT INFORMATION:
          NAME: Mistock, S. Leslie
          REGISTRATION NUMBER: 18,872
        REFERENCE/DOCKET NUMBER: 7683-021
        TELECOMMUNICATION INFORMATION:
          TELEPHONE: (212)790-9090
          TELEFAX: (212)869-8864/9741
          TELEX: 66141 PENNIE
        INFORMATION FOR SEQ ID NO: 2:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 6924 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: double
            TOPOLOGY: unknown
            MOLECULE TYPE: cDNA
            FEATURE:
              NAME/KEY: CDS
              LOCATION: 1..6924
US-08-015-973-2

Query Match           14.8%; Score 33; DB 1; Length 6924;
Best Local Similarity 49.7%; Pred.No. 0.25;
Matches   84; Conservative    0; Mismatches   85; Indels    0; Gaps    0.

QY      53 TCTTGATTCCTTACCCCTTAGAAGGCGCCTGAGTAATCACTCATTTCACTGAACAACA 112
         ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      5927 TCTTCATTTCATGATAACACTGCCTTAGGCCCATACTAGTAAGAAGCTGAGTCTGGACA 5986

QY      113 ATGGCTATGAAGGCACTTGTCGTTCGAANTGCACCCCAATGTGCCAGAGATGAACACTCA 172
         ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      5987 GTCATATTCATCCCTATGTTAAATGCACTCCATTCCTCTGAGACGAGCAAAACAAAGC 6046

QY      173 TTCACAATAATAAAGACATGTCAGCCAGCAGGCAATCTCTGTATCTGTTGA 221
         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      6047 TAGAATAAACATTCACGCTCCTGAGCCACTCAATATATACAGCAGACTGA 6095

RESULT      3
US-08-448-164-2
: Sequence 2, Application US/08448164
: Patent No. 5925536
: GENERAL INFORMATION:
: APPLICANT: Schlusser, Joseph
: TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
: TITLE OF INVENTION: PHOSPHATASE-BETA
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PENNIE & EDMONDS
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,164
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/015,973

```

[illegible]

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619-238-0999
 TELEFAX: 619-238-0062
 TELEX:
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1876 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 73..1581
 US-08-466-589-7

Query Match- 14.1%; Score 31.4; DB 2; Length 1876;
 Best Local Similarity 57.7%; Pred. No. 0.47;
 Matches 56; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 57 GATTCCTACCTCTGAGAGGGCCCTGATATTCATCATCAGTGAACAACATGG 116
 DB 366 GACTGCTTCCTTCCAGATGGCCAGATTGGAAACGACGATTCCTCTATAACAGTGC 425
 QY 117 CTATGAAGCATTGTGCTTGCATTCGACCCCATGTG 153
 DB 426 TGATGAGCGCTTGGAGCCGACATTCACACTAACGTG 462

RESULT 5

US-08-700-636-7
 Sequence 7, Application US/08700636
 Patent No. 5910582

GENERAL INFORMATION:

APPLICANT: Eliot, Kathryn J.
 APPLICANT: Ellis, Steven B.
 APPLICANT: Harpold, Michael M.
 TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESS: Pretty, Schroeder, Brueggemann & Clark
 STREET: 444 South Flower Street, Suite 2000
 CITY: Los Angeles
 STATE: CA
 COUNTRY: USA
 ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/700,636
 FILING DATE: 16-JUL-1996
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/028,031
 FILING DATE: 08-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Reiter, Stephen E.
 REGISTRATION NUMBER: 31,192
 REFERENCE/DOCKET NUMBER: P41 9368
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619-546-4737
 TELEFAX: 619-546-9392
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1876 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both

MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 73..1581
 US-08-700-636-7

Query Match- 14.1%; Score 31.4; DB 2; Length 1876;
 Best Local Similarity 57.7%; Pred. No. 0.47;
 Matches 56; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 57 GATTCCTACCTCTGAGAGGGCCCTGATATTCATCATCAGTGAACAACATGG 116
 DB 366 GACTGCTTCCTTCCAGATGGCCAGATTGGAAACGACGATTCCTCTATAACAGTGC 425
 QY 117 CTATGAAGCATTGTGCTTGCATTCGACCCCATGTG 153
 DB 426 TGATGAGCGCTTGGAGCCGACATTCACACTAACGTG 462

RESULT 6

US-08-467-574-7
 Sequence 7, Application US/08467574
 Patent No. 6022704

GENERAL INFORMATION:

APPLICANT: Eliot, Kathryn J.
 APPLICANT: Ellis, Steven B.
 APPLICANT: Harpold, Michael M.
 TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESS: Brown, Martin, Haller & McClain
 STREET: 1660 Union Street
 CITY: San Diego
 STATE: CA
 COUNTRY: USA
 ZIP: 92101-2926

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,574
 FILING DATE: June 5, 1995
 CLASSIFICATION: 336
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/028,031
 FILING DATE: March 8, 1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 6362-9949
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619-238-0999
 TELEFAX: 619-238-0062
 TELEX:

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1876 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 73..1581
 US-08-467-574-7

Query Match- 14.1%; Score 31.4; DB 3; Length 1876;
 Best Local Similarity 57.7%; Pred. No. 0.47;

Query Match	12.5%;	Score 27.8;	DB 3;	length 1438;
Best Local Similarity	59.5%;	Pred. NO. 6.5;		
Matches 47;	Conservative	0;	Mismatches 32;	Indels 0;

QY	57	GATCTTCACCTTCTMAAGGGGCCCTGAGTAATTCACATTCAGGTGMAACAACATGG	116
Db	891	GGCTCTTCACCTGCTGCAGGTCATTCAGGCTGAGTACTTATTCAGTCAATCACTTTGT	832
QY	117	CTATGAGGCATTCGCTT	135
Db	831	CCATCCAGAAAATGTCGGT	813

RESULT 10
US-09-499-884-11
; Sequence 11, Application US/09499884
; Patent No. 6265173

```

? TITLE OF INVENTION: DIAGNOSTIC TEST AND THERAPY FOR MANGANESE SUPEROXIDE DISMUTASES
? TITLE OF INVENTION: ASSOCIATED DISEASES
? FILE REFERENCE: 50229-180
? CURRENT APPLICATION NUMBER: US/09/499,864
? CURRENT FILING DATE: 2000-02-08
? NUMBER OF SEQ ID NOS: 12
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 11
? LENGTH: 3663
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-499-864-11

```

	Query Match	Best Local Similarity	Score 27.6;	DB 4;	Length 3663;
	Matches 45;	Conservative	0;	Mismatches 29;	Indels 0;
OY	150	TTGTCCGAAATGTAACACTCATCTACACAAATAAAGACATGTCGACCCAGCATCTCT	209		
Db	939	tgggcaacaataagaanaactcgttltctacaataataaaaaaalttagccagcagatgtr	998		
OY	210	GTATCTGTTTGAAG	223		
Db	999	gtatgcacctgtag	1012		

RESULT 11
 US-08-403-866-11
 Sequence 11, Application US/08403866
 Patent No 5643779
 GENERAL INFORMATION:
 APPLICANT: Ethlich, Stanislaw
 APPLICANT: Godon, Jean-Jacques
 APPLICANT: Renault, Pierre
 TITLE OF INVENTION: Nucleic acid coding for an alpha-acetoalactate
 TITLE OF INVENTION: synthase from Lactococcus and its applications
 NUMBER OF SEQUENCES: 16
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/403,866

ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30, 727

REFERENCE/DOCKET NUMBER: 20747/30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (716) 263-1600
 TELEFAX: (716) 263-1487
 TELE: 978450 (WOT)
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 12720 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: *Lactococcus lactis* subsp. *lactis*

Query	Match	Similarity	Score	DB	Length
Best	Local	Similarity	56.7%	Pred.	20
Matches	51	Conservative	0	Mismatches	39
				Indels	0
				Gaps	0
QY	69	TCTAGAGGGGCCCGATTAATTCACCTATTAGCTGGAACACCAATGGCTATGAAGGCAT	128		
DB	5296	TCAMGACAAATTCACATTTGATTTTACCCCATCAGCTAATCATCACAGCCTGTGACTT	5265		
QY	129	TGTCGTTGCATCGACCCCAATGTGCCAGA	158		
DB	5286	TCATTTTGAGATGACCCCATTTGGAAAGA	5295		

RESULT: 12
 5449756-10
 Patent No. 5449756
 APPLICANT: TANIGUCHI, TADATSUGU; HATAKEYAMA, MASANORI; MINAMOTO, SEIRO; KONO, TAKESHI; DOI, TAKESHI; MIYASAKA, MASAYUKI; TSUDO, MITSURU; KRAUSVYAMA, HAJIME
 TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2
 NUMBER OF SEQUENCES: 12
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/88,592
 FILING DATE: 9-JUL-1993
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 865,155
 FILING DATE: 08-APR-1992
 APPLICATION NUMBER: 487,059
 FILING DATE: 05-MAR-1990
 SEQ ID NO: 10:
 LENGTH: 1620
 5449756-10

Query Match	12.3%	Score 27.4	DB 6	Length 1620
Best Local Similarity	52.1%	Pred. No. 9.4		
Matches 61	Conservative 0	Mismatches 56	Indels 0	Gaps 0
106 AACACATGCGATGAAAGCATTTGTCGTTGCAATGACCCCAATGTCGACAGATGAA	165			
D0 1471 aatgacctcttgaggagcgccatgtgtccagaaggaacattctatgggaagaatcagagc	1530			
0Y : 167 ACACCTCATCAAAATTAAGAGCATGTCGACCCAGCATCTCTGATTTGTGAA	222			
D0 1531 agagagcagggcccaacctgtacacctgtgaacacggaatgtctatctgtcttcttaagaa	1587			

RESULT 13
5198359-3
Patent No. 5198359
APPLICANT: TANIGUCHI, TADATSUGU; JAPANESE, MASANORI;
MINAKO, O.; SEIJO, KONO, TAKEHI; DOI, TAKEHI; MIYASAKA, MASAYUKI
TSDO, MITSURU; KASUYAMA, NAJIME
TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2


```

:   NUMBER OF SEQUENCES: 9
:   CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/07/487,059
:   FILING DATE: 05-MAR-1990
:   SEQ ID NO:3:
:   LENGTH: 2306
5198359-3

```

```

Query Match      12.3%; Score 27.4; DB 6; Length 2306;
Best Local Similarity 52.1%; Pred. No. 11;
Matches 61; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

```

```

QY 106 AACAAATGCGTATGAAGGCAATGTGCTGCAATCGACCCCAATGTCGACAAAGATGAA 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1845 aatagtctcgggagcagcagctgtcccaagaagcaacctcattcggcaagatcagagc 1904
QY 166 ACACCTCATTCACAAATAAAGACATGTCGACCCAGGCACTCTGATCTGTTTGA 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1905 agagggcaggcccatcctgacctgaacacgatgctcattcttcttaagaa 1961

```

```

RESULT 14
5449756-3
:   Patent No. 5449756
:   APPLICANT: TANIGUCHI, TADATSUGU; HATAKEYAMA, MASANORI; MINAMOTO,
:   SEJIRO; KONO, TAKESHI; DOI, TAKESHI; MIYASAKA, MASAYUKI; TSUDO,
:   MITSURU; KANASUTAMA, HAJIME
:   TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2
:   NUMBER OF SEQUENCES: 12
:   CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/88,592
:   FILING DATE: 9-JUL-1993
:   PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 865,155
:   FILING DATE: 08-APR-1992
:   APPLICATION NUMBER: 487,059
:   FILING DATE: 05-MAR-1990
:   SEQ ID NO:3:
:   LENGTH: 2306
5449756-3

```

```

Query Match      12.3%; Score 27.4; DB 6; Length 2306;
Best Local Similarity 52.1%; Pred. No. 11;
Matches 61; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

```

```

QY 106 AACCAATGCGTATGAAGGCAATGTGCTGCAATCGACCCCAATGTCGACAAAGATGAA 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1845 aatagtctcgggagcagcagctgtcccaagaagcaacctcattcggcaagatcagagc 1904
QY 166 ACACCTCATTCACAAATAAAGACATGTCGACCCAGGCACTCTGATCTGTTTGA 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1905 agagggcaggcccatcctgacctgaacacgatgctcattcttcttaagaa 1961

```

```

RESULT 15
US-08-841-349-3
:   Sequence 3, Application US/08841349B
:   Patent No. 5955594
:   GENERAL INFORMATION:
:   APPLICANT: MISHRA, LOPA
:   TITLE OF INVENTION: GENES CODING PROTEINS FOR EARLY LIVER DEVELOPMENT.
:   FILE REFERENCE: XX/PO4470050
:   CURRENT APPLICATION NUMBER: US/08/841,349B
:   CURRENT FILING DATE: 1997-04-30
:   NUMBER OF SEQ ID NOS: 18
:   SOFTWARE: Patentln Ver. 2.0
:   SEQ ID NO 3
:   LENGTH: 6960
:   TYPE: DNA
:   ORGANISM: Mus musculus
:   FEATURE:

```

```

:   NAME/KEY: CDS
:   LOCATION: (333)..(6794)
US-08-841-349-3

```

```

Query Match      12.2%; Score 27.2; DB 2; Length 6960;
Best Local Similarity 55.2%; Pred. No. 21;
Matches 53; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

```

QY 103 CTGAACACCAATGCGCTATGAAGGCAATGTGCTGCAATCGACCCCAATGTCGACAAAGAT 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3846 cttaataccaggaggtatgttttggctcattcgaatgcccaccaccctggaaagagct 3905
QY 163 GAACACTCATTCACAAATAAAGACATGTCGACCC 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3906 gaagcagccattaaagcaggaggaattcattgacc 3941

```

```

Search completed: April 3, 2002, 20:36:09
Job time: 40727 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 10:34:08 ; Search time 1321.64 Seconds
(Without alignments)
144.656 Million cell updates/sec

Title: US-09-049-696-1
Perfect score: 223
Sequence: 1 GAATACACAGGAGATGTAC.....ATCTCTGATCTGTTGAAG 223

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: N.Geneseq_1101.*
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	223	100.0	2854	22	AAH34879
2	223	100.0	2854	22	AAH34879
3	223	100.0	2854	22	AAH34879
4	223	100.0	2854	22	AAH34879
5	223	100.0	2854	22	AAH34879
6	221.4	99.3	3311	20	AAH09840
7	221.4	99.3	3311	20	AAH09840
8	199	89.2	2742	22	AAH46102
9	138.2	62.0	2739	22	AAH46101
10	138.2	62.0	2843	22	AAH46120
11	138.2	62.0	2931	20	AAH81925

Result No.	Score	Query Match	Length	DB ID	Description
12	115	51.6	401	21	AAH7755
13	115	51.6	401	21	AAH7755
14	100.6	45.1	3265	21	AAH28493
15	100.6	45.1	3265	21	AAH28493
16	100.6	45.1	3265	21	AAH28493
17	99.9	44.4	375	22	AAH44241
18	91.2	40.9	350	22	AAH72944
19	91.2	40.9	350	22	AAH72944
20	91.2	40.9	350	22	AAH72944
21	68	30.5	454	22	AAH69537
22	68	30.5	454	22	AAH69537
23	50.8	22.8	332	21	AAH03649
24	50.8	22.8	332	21	AAH03649
25	50.8	22.8	332	21	AAH03649
26	50.8	22.8	332	21	AAH03649
27	50.8	22.8	332	20	AAH22657
28	50.8	22.8	332	21	AAH65886
29	50.8	22.8	332	21	AAH65886
30	50.8	22.8	332	21	AAH65886
31	50.8	22.8	332	21	AAH65886
32	49.2	22.1	310	20	AAH81926
33	49.2	22.1	310	20	AAH81926
34	33.4	15.2	304	22	AAH66121
35	33.4	15.0	349	22	AAH68527
36	33.4	14.8	6924	15	AAH67283
37	33.4	14.8	8146	22	AAH73240
38	32	14.3	5524	10	AAH90123
39	31.8	14.3	8930	19	AAH22834
40	31.4	14.1	1416	22	AAH90382
41	31.4	14.1	1509	21	AAH58395
42	31.4	14.1	1509	22	AAH90380
43	31.4	14.1	1509	22	AAH90385
44	31.4	14.1	1509	22	AAH90386
45	31.4	14.1	1509	22	AAH90387

ALIGNMENTS

RESULT: 1
ID: AAH34879 standard; CDNA; 2854 BP.
AC: AAH34879;
XX: 73-SEP-2001 (first entry)
CI: Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
DE: Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
XX: Human colon cancer; colon cancer antigen; diagnosis; detection;
KW: colorectal carcinoma; chromosome 1; ss.
XX: Homo sapiens.
OS: Homo sapiens.
PN: W0200122920-A2.
XX: 95-APR-2001.
PO: 26-SEP-2000; 2000MO-US26524.
PF: 20-SEP-1999; 9905-0157137.
XX: 03-NOV-1999; 9905-0163280.
PR: (HUMA-) HUMAN GENOME SCI INC.
FA: Ruben SM, Barash SC, Birse CE, Rosen CA;
XX: WPI, 2001-235357/24.
DR: P. PSDB; AAG75174.
XX: Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT: useful for preventing, diagnosing and/or treating colorectal cancers -
XX: Mutant human alpha

PS Claim 1: Page 3462-3463; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB7789 represent sequences used in the exemplification of the
CC present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

CC XX
SQ Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match. 100.0%; Score 223; DB 22; Length 2854;
Best Local Similarity 100.0%; Pred. No. 4e-63;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAATCACAAGGAGATGATGACGATGAGGCCATTAAAGCTGTGTTCACTTGATT 60
DB 11 gaattacacaggagatgacgaatgagggccattaaagctgtgttcattcgtgatt 70
OY 61 CTTCACCTTCTGAAGGGGCCCTGAGTAATTCATTCAGCTGAAACAATGGCTAT 120
DB 71 ctccacctctgaaggagggccctgagtaattcctcattcagctgacacaaatgctat 130
OY 121 GAAGCATTTGCTGTCATTCGACCCCAATGTCGAGAAAGATGAACACTTCATTAACA 180
DB 131 gaagcatctgtctgtgacatgcacccaatgtgcagaaagatgaacactctcattcaaca 190
OY 181 ATTAAGACATGGGTGACCCAGCAGCATCTGTATCTGTTGAAG 223
DB 191 ataaagacatggtgacccagcagcatctgtatctgtttgaag 233

RESULT 2
AAH31787 standard; cDNA: 2854 BP.

AC AAF81787;
XX
DT 12-JUN-2001 (first entry)
XX
DE Human secreted protein gene 1 SEQ ID NO:11.
XX
XX Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
KW dermatological; immunosuppressive; antineoplastic; anti-HIV;
KW immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;
KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
KW antialzheimer; antiparkinsonian; antimicrobial; vulnary; gene therapy;
KW immune disorder; hyperproliferative disorder; cardiovascular disease;
KW cancer; angiogenic disorder; neurological disorder; infectious disease;
KW wound healing; regeneration; chemotaxis; chromosome 1; ss.
XX
OS Homo sapiens.
XX
PN WO200112775-A2.
XX
PD 22-FEB-2001.
XX
PF 16-AUG-2000; 2000MO-US22325.
XX

PR 17-AUG-1999; 99US-0149182.
XX
XX (HOMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;
PI Blase CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;
DR WPI; 2001-147550/15.
XX P-PsDB; AAB74733.
XX
XX Claim 1: Page 441; 485pp; English.

PT Nucleic acids encoding 25 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy.

PS
XX
XX AAF81787 to AAF81817 encode the human secreted proteins given in AAB74733
CC to AAB74772. Human secreted proteins can have activities based on the
CC tissues and cells they are expressed in. Example of activities include:
CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;
CC antineoplastic; anti-HIV; immunostimulant; cytostatic; cardiant;
CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
CC anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and
CC vulnary. Human secreted proteins can be used in gene therapy and
CC vaccine. Human secreted proteins nucleotide sequences (NMI) and proteins
CC (PPI) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. For example, NMI
CC and PPI may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of proteins by expressing inactive proteins or
CC to supplement the patient's own production of polypeptides. Disorders
CC may be prevented, diagnosed and/or treated include immune disorders,
CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
CC angiogenic disorders, neurological disorders, infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to
CC AAF81786 and AAB74732 represent sequences used in the exemplification of
CC the present invention.

XX
SQ Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match. 100.0%; Score 223; DB 22; Length 2854;
Best Local Similarity 100.0%; Pred. No. 4e-63;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAATCACAAGGAGATGATGACGATGAGGCCATTAAAGCTGTGTTCACTTGATT 60
DB 11 gaattacacaggagatgacgaatgagggccattaaagctgtgttcattcgtgatt 70
OY 61 CTTCACCTTCTGAAGGGGCCCTGAGTAATTCATTCAGCTGAAACAATGGCTAT 120
DB 71 ctccacctctgaaggagggccctgagtaattcctcattcagctgacacaaatgctat 130
OY 121 GAAGCATTTGCTGTCATTCGACCCCAATGTCGAGAAAGATGAACACTTCATTAACA 180
DB 131 gaagcatctgtctgtgacatgcacccaatgtgcagaaagatgaacactctcattcaaca 190
OY 181 ATTAAGACATGGGTGACCCAGCAGCATCTGTATCTGTTGAAG 223
DB 191 ataaagacatggtgacccagcagcatctgtatctgtttgaag 233

RESULT 3
AAH33285 standard; cDNA: 2867 BP.

AC AAF81787;
XX
XX
XX 03-SEP-2001 (first entry)
XX
XX Human colon cancer antigen encoding cDNA SEQ ID NO:341.
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW

KW colorectal carcinoma; chromosome 1; ss.
 XX Homo sapiens.
 XX MO200122920-A2.
 PN 05-APR-2001.
 PD 28-SEP-2000; 2000MO-US26524.
 PF 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI WPI; 2001-235357/24.
 DR P-PSDB; AAG3854.
 DR Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX Claim 1; Page 2452-2453; 9803pp; English.
 PS AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated PS,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other:
 SQ
 Query Match 100.0%; Score 223; DB 22; Length 2867;
 Best Local Similarity 100.0%; Pred. No. 4e-63;
 Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAATATCAGGAGATGTACAGCAATGGGCGCATTTAAGAGTCTGTGTCATCTTATT 60
 DB 14 gaatcacagggagatgtacagcaatgggcccatttaagagttctgtgtcattctgatt 73
 QY 61 CTTCACCTTTAGAGGGGCGCTGAGTAATTCATCTCATTGAGTGAACAAATGGCTAT 120
 DB 74 ctccacctctagaagggccctgagtaattcactcattcagctgaacaacaatgctat 133
 QY 121 GAAGGCAATTCGTTGCAATCGAACCCCAATGTGCCAGACAGATGAACACTTCATCAACAA 180
 DB 134 gaaggaatctgctgttgaatcgaaacccaatgtgccagaagatgaacaacatctcaacaa 193
 QY 181 ATAAAGACATGATGATGCCAGCATCTGTATCTGTTTAAG 223
 DB 194 ataaagacatggtgacccagcatctctgtatctgttgaag 236

AC AA209840;
 XX 26-NOV-1999 (first entry)
 DE Human membrane spanning protein MSP-5 CDNA fragment 2.
 XX Mammalian spanning protein; MSP; human; treatment; diagnosis; prevention;
 KW neoplastic disorder; immunological disorder; reproductive disorder;
 XX MSP-5; ds.
 OS Homo sapiens.
 XX MO9946380-A2.
 PN 16-SEP-1999.
 PD 09-MAR-1999; 99MO-US05073.
 PF 13-MAR-1998; 98US-0039064.
 PR (INCY-) INCYTE PHARM INC.
 XX Tang YT, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;
 PI Giegler KJ, Kaser MR, Baughn MR, Shah P;
 PT WPI; 1999-551409/46.
 DR P-PSDB; AAY33298.
 DR New human membrane spanning proteins used to, e.g. prevent and treat
 PT neoplastic disorders -
 PS Example 1; Page 80-81; 81pp; English.
 XX This invention describes novel human membrane spanning proteins (MSPs),
 CC and the polynucleotides encoding them. The products of the invention are
 CC used to diagnose, prevent and treat neoplastic, immunological and
 CC reproductive disorders. This sequence encodes a human membrane spanning
 CC protein MSP-5 fragment.
 XX Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other:
 SQ
 Query Match 100.0%; Score 223; DB 20; Length 3111;
 Best Local Similarity 100.0%; Pred. No. 4.2e-63;
 Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAATATCAGGAGATGTACAGCAATGGGCGCATTTAAGAGTCTGTGTCATCTTATT 60
 DB 30 gaatcacagggagatgtacagcaatgggcccatttaagagttctgtgtcattctgatt 69
 QY 61 CTTCACCTTTAGAGGGGCGCTGAGTAATTCATCTCATTGAGTGAACAAATGGCTAT 120
 DB 70 ctccacctctagaagggccctgagtaattcactcattcagctgaacaacaatgctat 129
 QY 121 GAAGGCAATTCGTTGCAATCGAACCCCAATGTGCCAGACAGATGAACACTTCATCAACAA 180
 DB 130 gaaggaatctgctgttgaatcgaaacccaatgtgccagaagatgaacaacatctcaacaa 189
 QY 181 ATAAAGACATGATGATGCCAGCATCTGTATCTGTTTAAG 223
 DB 290 ataaagacatggtgacccagcatctctgtatctgttgaag 232
 RESULT 5
 ID AA129502 standard; CDNA; 3311 BP.
 AC AA129502;
 XX 12-OCT-2001 (first entry)
 DE 3502P determined CDNA sequence.

KW	Human: Immunotherapy; diagnosis; colon cancer; colon tumour;
KW	Immunogenic; gene therapy; vaccine; colonic cancer; ss.
OS	Homo sapiens.
PN	WO200149716-A2.
PN	12-JUL-2001.
XX	
XX	
XX	29-DEC-2000; 2000WO-US35596.
XX	
XX	30-DEC-1999; 90US-0476296.
XX	10-JAN-2000; 2000US-0480321.
PR	15-FEB-2000; 2000US-0504629.
PR	06-MAR-2000; 2000US-0519444.
PR	19-MAY-2000; 2000US-0575251.
PR	29-JUN-2000; 2000US-0609448.
XX	28-AUG-2000; 2000US-0649811.
PA	(CORI-) CORIXA CORP.
PI	
PI	Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;
PI	King GE, Wang T, Jiang Y;
XX	
XX	WPI: 2001-441847/47.
XX	
XX	Colon tumor associated proteins and nucleic acids useful for the
PT	prevention, diagnosis and treatment of colonic cancer -
PT	
PS	Claim 2: Page 425-426; 472pp: English.
XX	
XX	The present invention describes colon tumour associated proteins (I) and
CC	the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC	(I) and (II) can be used in gene therapy and vaccine production. (I) and
CC	(II) may be used in the prevention, diagnosis and treatment of diseases
CC	associated with inappropriate colon tumour associated protein (TCAP)
CC	expression, such as colonic cancer. For example, (I) and (II) may be
CC	used to treat disorders associated with decreased expression by
CC	rectifying mutations or deletions in a patient's genome that affect the
CC	activity of TCAPs by expressing inactive proteins or to supplement the
CC	patients own production of them. Additionally, (II) may be used to
CC	produce the TCAP proteins, by inserting the nucleic acids into a host
CC	cell culturing the cell to express the protein. (II) and its
CC	complementary sequences may also be used as DNA probes in diagnostic
CC	polymerase chain reaction (PCR) and hybridisation assays to detect and
CC	quantitate the presence of similar nucleic acids in samples, and
CC	therefore which patients may be in need of restorative therapy. (I) may
CC	also be used as antigens in the production of antibodies against TCAPs
CC	and in assays to identify modulators of TCAP expression and activity.
CC	Anti-(I) antibodies and antagonists may also be used to down regulate
CC	TCAP expression and activity. The anti-(I) antibodies may also be used
CC	as diagnostic agents for detecting the presence of TCAPs in samples
CC	(e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
CC	and AA24494 to AA24523 represent nucleotide and amino acid sequences
XX	given in the exemplification of the present invention.
XX	
XX	Sequence 3311 BP: 1028 A; 692 C; 742 G; 849 T; 0 other;
QY	
QY	Query Match 99.3%; Score 221.4; DB 22; Length 3311;
QY	Best Local Similarity 99.6%; Pred. No. 1.4e-62;
QY	Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0
Db	
Db	1 GAAATCACAGGAGGAGNTGTCAGCAATGGGGCCATTAAAGAGTCTGTGTCATCTTGATT 60
Db	
Db	328 ggaatcacagggagagatgacagcaatggggccatttaagaagtctgtatcattcttgatt 387
QY	
QY	61 CTTCACCTTCTGAAGAGGGCCCTGAGTATTCATCTATTCAGCTGGAACAACATGGCTAT 120
Db	
Db	388 cttaacctctctagaaggccctgtgaattcattcattcagctgtaacaacaatgcttat 447
QY	
QY	121 GAAGCATTTGCTTGCAATGCACCCCAATGTGCCAGAAATGAACACTCATTCACAAC 180
QY	

```

Db      448 gaaagcattgctgttgcataatgcaccccaatgylgcagaagtgaataaccattcaaca 507
QY      181 ATPAAGCACAATGCTGCACCCAGCATCTCTGTATCTGTTGAAG 223
        |||||||
Db      508 ataagaacattgttgaccaggcatctctgtacttgtttaag 550
        |||||||

RESULT# 6
AAH46124
ID. AAH46124 standard; cDNA; 2825 BP.
AC AAH46124;
DE Human CLCA1 cDNA, SEQ ID NO:26.
KW Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
KV expression inhibition; antisense therapy; gene therapy;
KW chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
.. ..
OS Homo sapiens.
FX Key Location/Qualifiers
FT CDS 23..2767
    /tag= a
    /product= "Human CLCA1"
    /transl_except= (pos:476..478, aa:Lys)
FT FT
FT FT
PN WO200138530-AI.
PX 31-MAY-2001.
PY 22-NOV-2000; 2000WO-JP08232.
PR 24-NOV-1999; 99JP-0333479.
PR 27-APR-2000; 2000JP-0127589.
PA (TAKE ) TAKEDA CHEM IND LTD.
PX Naikanishi A, Morita S;
DR WPI: 2001-355935/37.
PS P-SDB; AAB73716.
PX New antisense nucleotide, useful for treatment and prevention of
PX bronchial asthma and chronic obstructive pulmonary disease -
Example 5; Page 92-94; 104pp; Japanese.

The invention relates to an antisense nucleotide targeted to the mouse
Gov-5 gene (coding sequence shown in AAH46101) or its human counterpart,
the CLCA1 gene (coding sequence shown in AAH46102). The invention also
relates to an antibody specific for the Gob-5 protein, medical and
diagnostic compositions containing the antisense nucleotide or the
antibody, and methods and kits for screening for compounds which inhibit
the protein. Gov-5 and CLCA1 are proteins expressed by goblet cells.
The antisense-polynucleotides and antibody are therefore useful for the
treatment and prevention of bronchial asthma and chronic obstructive
pulmonary disease. The present sequence represents human CLCA1 cDNA.

Sequence 2825;BP; 864 A; 626 C; 637 G; 698 T; 0 other;

Query Match          99.1%; score 221; DB 22; Length 2825;
Best Local Similarity 100.0%; Pred. No. 1,8e-62;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3 AATCACAGCAGAGATTACAGCAATGGGCCAATTAAGAGTTCCTGTTCAATCTGATTCCT 62
|||||
db 1 aatcacagagagattacagcaatgggccaatttaagagttctctgttcattcttatctt 60

```


KW Immunotherapy; diagnosis; progression; ss.
XX Homo sapiens.
OS
XX MO200037643-A2.
XX
XX 29-JUN-2000.
PD
XX 23-DEC-1999; 99MO-US30909.
XX
XX 23-DEC-1998; 98US-0221298.
PR 02-JUL-1999; 99US-0347496.
PR 22-SEP-1999; 99US-0401064.
PR 19-NOV-1999; 99US-0444242.
PR 02-DEC-1999; 99US-0454150.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk J;
PI Wang T, Yugu J;
XX
XX WPI: 2000-442671/38.
XX
XX New colon tumor polypeptides used to inhibit the development of cancer,
PT especially colon cancer, and for diagnosing and monitoring the
PT progression of the cancer -
XX
XX Claim 1; Page 97; 229pp; English.
XX
XX Sequences AAA7722-A78199 represent 478 cDNAs encoding proteins or
CC portions of proteins which are associated with human colon tumours.
CC The invention also specifically discloses 8 human colon tumour proteins
CC (AA11887-811904). The nucleic acids, the polypeptides they encode, and
CC antigen presenting cells (APCs), preferably dendritic cells) expressing
CC such polypeptides may be used in vaccines that target tumour cells,
CC especially colon tumour cells, thereby inhibiting the development of
CC cancer. T-cells specific for the polypeptide expressed by the APC are
CC used to remove tumour cells from biological samples, especially blood or
CC fractions thereof. The sample or the isolated T-cells specific for the
CC polypeptide can then be used to inhibit cancer development. CD4+ and/or
CC CD8+ T-cells from a patient may be incubated with a polypeptide or
CC nucleic acid of the invention, or an APC expressing such a polypeptide,
CC to cause the proliferation of specific T-cells. The T-cells can be
CC cloned and then administered back to the patient to inhibit cancer
CC development. Nucleic acids encoding the polypeptides and antibodies
CC against the polypeptides may be used to determine the expression level
CC of a tumour protein of the invention, and therefore to determine whether
CC cancer cells are present. Such diagnostic methods may also be used to
CC monitor the progression of a cancer by repeating the processes at time
CC intervals, and comparing the current result to previous results. The
CC present sequence represents a cDNA encoding a human colon tumour
CC polypeptide.
XX
XX Sequence 401 BP; 127 A; 80 C; 95 G; 99 T; 0 other;
XX
XX
XX Query Match 51.6%; Score 115; DB 21; Length 401;
XX Best Local Similarity 100.0%; Pred. No. 4.5e-28;
XX Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 109 AACATGCTATGAGCATGTCGTTGCAATGACCCCAATGCGCAAGATGAACA 168
DB 1 aacatgctatgaagcatgtctgcaatgcagcccaatgctcagaaatgaaca 60
XX
XX 169 CTCATTCAACAATAAGAGCATGTGACCCAGCATCTCTGTATCTGTTGAAG 223
DB 61 ctcattcaacaataaagagcatgtgacccagcatctctgtatctgttgaag 115
XX
XX
XX RESULT 13
XX AA128493
XX ID AA128493 standard; cDNA; 401 BP.
XX

AC AA128493;
XX
XX 12-OCT-2001; (first entry)
DE Colon tumour related cDNA sequence Contig 32.
XX
XX Human; Immunotherapy; diagnosis; colon cancer; colon tumour;
KW Immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX
XX Homo sapiens.
OS
XX MO200149716-A2.
XX
XX 29-DEC-2000; 2000MO-US35596.
XX
XX 30-DEC-1999; 99US-0476296.
PR 30-JAN-2000; 2000US-0480321.
PR 15-FEB-2000; 2000US-0504629.
PR 06-MAR-2000; 2000US-0519444.
PR 19-MAY-2000; 2000US-0575251.
PR 29-JUN-2000; 2000US-0609448.
PR 26-AUG-2000; 2000US-0649811.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX
XX WPI: 2001-441847/47.
XX
XX Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer -
PT
XX
XX Claim 2; Page 139; 472pp; English.
XX
XX The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC patients own production of them. Additionally, (II) may be used to
CC produce the TCAP proteins, by inserting the nucleic acids into a host
CC cell culturing the cell to express the protein. (II) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC polymerase chain reaction (PCR) and hybridisation assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. (I) may
CC also be used as antigens in the production of antibodies against TCAPs
CC and in assays to identify modulators of TCAP expression and activity.
CC Anti-(I) antibodies and antagonists may also be used to down regulate
CC TCAP expression and activity. The anti-(I) antibodies may also be used
CC as diagnostic agents for detecting the presence of TCAPs in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
CC and AA124494 to AA124523 represent nucleotide and amino acid sequences
CC given in the exemplification of the present invention.
XX
XX Sequence 401 BP; 127 A; 80 C; 95 G; 99 T; 0 other;
XX
XX
XX Query Match 51.6%; Score 115; DB 22; Length 401;
XX Best Local Similarity 100.0%; Pred. No. 4.5e-28;
XX Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 109 AACATGCTATGAGCATGTCGTTGCAATGACCCCAATGCGCAAGATGAACA 168
DB 1 aacatgctatgaagcatgtctgcaatgcagcccaatgctcagaaatgaaca 60
XX

OY 169 CTCATTCACAATTAAGACATGTCACCAAGCATCTCTGTATCTTTGAG 223
DB 61 ctcattcaacaataaagacatgtagaccagcgcattctgtatctgttgaag 115

RESULT 14
AAZ65095
AAZ65095 standard; CDNA; 3265 BP.

XX
AC AAZ65095;
XX
DT 05-APR-2000 (first entry)
XX
DE Membrane-bound protein PRO1124 encoding CDNA.
XX
KM Membrane-bound polypeptide: PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping; ss.
XX
OS Homo sapiens.
XX
PN W09963088-A2.
XX
PD 09-DEC-1999.
XX
PF 02-JUN-1999; 99WO-US12252.
XX
PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 03-JUN-1998; 98US-0087759.
PR 04-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.

PR 10-JUN-1998; 98US-0089948.
PR 10-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 22-JUN-1998; 98US-0090349.
PR 22-JUN-1998; 98US-0090355.
PR 22-JUN-1998; 98US-0090429.
PR 22-JUN-1998; 98US-0090431.
PR 22-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 26-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091476.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 02-JUL-1998; 98US-0091978.
PR 02-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 30-JUL-1998; 98US-0093339.
PR 04-AUG-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.

Thu Apr 4 09:27:22 2002

us-09-049-696-1.rng

Page 11

Db 198 agctctacgtacactgttgaag 220

Search completed: April 4, 2002, 03:17:07
Job time: 60179 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 10:15:18 ; Search time 22700.8 Seconds

(without alignments)
162.059 Million cell updates/sec

Title: US-09-049-696-1

Perfect score: 223
Sequence: 1 GAAATCAGGAGAGATGTAC.....ATCTCTGATCTGTTGAAG 223

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: 1: gb_ba: 2: gb_hg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_om: 20: em_or: 21: em_ov: 22: em_pat: 23: em_ph: 24: em_pl: 25: em_ro: 26: em_sts: 27: em_sy: 28: em_un: 29: em_vl: 30: em_hgo_hum: 31: em_hgo_inv: 32: em_hgo_rod: 33: em_hgo_hum: 34: em_hgo_inv: 35: em_hgo_rod: 36: em_hgo_other:

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	221.4	99.3	3311	6	AX193489 Sequence
2	221.4	99.3	3311	9	AF039400 Homo sapi
3	203	91.0	2826	9	AF127036 Homo sapi
4	187	33.9	113764	9	HS1651E10
5	187	33.9	124096	2	AL122002 Human DNA
6	187	33.9	164891	2	AL135498 Homo sapi
7	185.4	33.1	35278	2	AL358950 Homo sapi
8	165.4	74.2	3079	4	AF095584 Sus scrofa
9	138.2	52.0	2937	10	AB017156
10	115	51.6	401	6	AX192467 Sequence
11	100.6	45.1	3204	9	AF127035 Homo sapi
12	100.6	45.1	3221	9	AK000072
13	100.6	45.1	3265	6	AX092338 Sequence
14	99	44.4	375	6	AX071171
15	91.2	40.9	350	6	AX188011 Sequence
16	91.2	40.9	454	6	AX185116 Sequence
17	91	40.8	1265	4	AF001263 Bos tauru
18	91	40.8	3317	4	AF001261 Bos tauru
19	89.4	40.1	2984	4	BT036445
20	86.6	38.6	3288	4	AF001262
21	85.2	38.2	3415	9	AF043976
22	85.2	38.2	2765	10	AF115852
23	73.8	33.1	3058	10	AF108501
24	73.8	33.1	3137	10	BC008147
25	73.8	32.6	175591	2	AC068071
26	72.8	32.4	3022	10	AF047838
27	72.2	32.4	3471	10	AF052746
28	72.2	32.4	140718	2	AL356270
29	71.2	31.9	227	6	AX186451
30	68	30.5	2832	9	AF127980
31	50.8	27.8	2970	9	AF043977
32	50.8	27.8	4077	6	AX054697 Sequence
33	50.8	27.8	3604	9	AB026833
34	49.2	22.1	330743	14	PM042580
35	40.6	13.2	130369	9	AC074090
36	37.8	17.0	177298	2	AC068696
37	37.8	16.2	110000	2	AC018855
38	36.8	15.5	177578	2	LMFCHR34_04
39	36.2	15.2	151524	2	AC019321
40	35.8	15.1	151524	9	AC006060
41	35.4	15.9	135039	2	AC018784
42	35.2	15.8	108916	2	AC015593
43	35	15.7	156998	2	AP001644
44	35	15.7	200798	2	AP001644
45	34.8	15.6	138517	9	AC019288

ALIGNMENTS

RESULT 1
LOCUS AX193489 3311 bp DNA
DEFINITION Sequence 1056 from Patent WO0149716.
ACCESSION AX193489
VERSION AX193489.1 GI:15211440
KEYWORDS
SOURCE
ORGANISM Human
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Xu, J., Lodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J.,
Stolicki, A., King, G.E., Wang, T. and Jiang, Y.
Comounds for immunotherapy and diagnosis of colon cancer and
metabls for their use
Patent: WO 0149716-A 1056 12-JUL-2001;
CORONA CORPORATION (US)

AUTHORS
TITL:
JOURNAL
FEATURES
SOURCE

1. 3311
Location/Qualifiers

/product="calcium activated chloride channel protein 1"
 /protein_id="AAD25487.1"
 /db_xref="GI:4585469"
 /translation="MGKSSVFLITLHLLEGALSLSLIOLANNNGEYIVADPNP
 EDPTLQOIKIDMTQASILYLPFRATGRFRFYKNAVILIPETWTKADYVPRKLETTKNA
 DVLVESTPGNDPEYEQMGNGCKGKERIHILPDIATKRLAEYGPQRAFEHNAH
 LMGVFEVDNDEKPEYLSNGRLOAVRCSAGITGTNNKCOGSGCTKCTKCTFVGL
 YKGECEVQSROTERKASIMFAQHVDSIVECTEQNHKEAPKQKONKLSRTWVLI
 RQSEDKKTPMTOPNPPTFSILOIGRIVCLIDKSGSMATNGNINLNNOGOLF
 LQPELVSGWGVTPEDSAHNSSELIOINSGSDRDLAKRIPAAAGSGRISIGELISA
 FTVIRKRYPTDSEIVLITDGEDNTISGCFNEKOSGATIHVALPSSAOLEBELSK
 MTGGLQTVASIDVONNGLIDAFGALSNGAVSORSIOLESKGLTQNSQNMWNTIV
 DSTVGKDTFLITWTQPOIILMDPSGOKGFFVNDKTKMAVLOIPGIATKVKY
 SIQASSQTLITVTSSASNATLPIVTSKTKNDTSKPSPLVYVNIROGASPIIRA
 SVTALIESVNGKVTLEILDNGAGADATDDGVSRFTYDIDNGRSVYKVRALGGV
 AARRVYPOQSGALYTPWIEDEIOMNPREPINNDVDVQHOVCSSRISGSPAS
 DYPNAPIDLPFGQITDLKAEIHGCSLNLWTAGDDTDHGTAKVILIRISTSLD
 LDKFNESLQVNTTALIPKANSSEVFLKPEPTENCTDLFIQAVDKDLKSEI
 SNIRVSLFIPPTPEPTSPETSAPCNHINISIPGILIMKIMKIGELQSLISA"

BASE COUNT 875 a 623 c 632 g 696 t
 ORIGIN
 Query Match 91.0%; Score 203; DB 9; Length 2826;
 Best local Similarity 100.0%; Pred. No. 1.8e-51;
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 AGCAATGGGGCCATTAGAGTTCTGTTCATCTGATCTTACCTTTAGAGGGC 80
 |||||||
 DB 1 AGCAATGGGGCCATTAGAGTTCTGTTCATCTGATCTTACCTTTAGAGGGC 60
 |||||||
 QY 81 CCGAGTAATTCATCTAGTCAGCAACAAAGGCTATGAGGAGCATTCGTCGAAT 140
 |||||||
 DB 61 CCGAGTAATTCATCTAGTCAGCAACAAAGGCTATGAGGAGCATTCGTCGAAT 120
 |||||||
 QY 141 CGACCCCAATGTGCGAGAAAGATGAACATCTATTCACAAATTAAGACATGTGACCA 200
 |||||||
 DB 121 CGACCCCAATGTGCGAGAAAGATGAACATCTATTCACAAATTAAGACATGTGACCA 180
 |||||||
 QY 201 GGCATCTCTATCTGTTTGAAG 223
 |||||||
 DB 181 GGCATCTCTATCTGTTTGAAG 203
 |||||||

RESULT 4
 HSJ651E10 113764 bp DNA PRI 27-MAY-2000
 LOCUS Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1,
 complete sequence.
 ACCESSION AL122002
 VERSION AL122002.16 GI:8247274
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 113764)
 Brown, A.
 Direct Submision
 Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jun 4, 2000 this sequence version replaced gi:8247030.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submision
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence has been finished according to sequence map criteria
 as follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu), where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr1
 RP4-651E10 is from the library RP4-4 constructed at the Roswell
 Park Cancer Institute by the group of Pieter de Jong. For further
 details see http://baepac.med.buffalo.edu/
 VECTOR: pCIRAC2
 This sequence is the entire insert of clone RP4-651E10.
 Location/Qualifiers
 1. 113764
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="p22.3-31.1"
 /clone="RP4-651E10"
 /clone_11b="RP4-4"

BASE COUNT 34890 a 21989 c 22351 g 34534 t
 ORIGIN

Query Match 83.9%; Score 187; DB 9; Length 113764;
 Best local Similarity 100.0%; Pred. No. 2.4e-46;
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAATCAGAGGAGATGTACAGCAATGGGCCATTTAAGACTTGTGTTTCATT 60
 |||||||
 DB 1918 GAAATCAGAGGAGATGTACAGCAATGGGCCATTTAAGACTTGTGTTTCATT 1977
 |||||||
 QY 61 CTTACCTTCTAGAGGGCCCTGAGTAATTCATCTAGTCAGTCAGCAACATGGCTAT 120
 |||||||
 DB 1978 CTTACCTTCTAGAGGGCCCTGAGTAATTCATCTAGTCAGTCAGCAACATGGCTAT 2037
 |||||||
 QY 121 GAAGCATGTGCTGTAATGACCCCAATGTGCCAGAAAGATGAACATCTATTCACAA 180
 |||||||
 DB 2038 GAAGCATGTGCTGTAATGACCCCAATGTGCCAGAAAGATGAACATCTATTCACAA 2097
 |||||||
 QY 181 ATTAAGC 187
 |||||||
 DB 2098 ATTAAGC 2104
 |||||||

RESULT 5
 AL354988
 LOCUS AL354988 124096 bp DNA HTG 23-JAN-2001
 DEFINITION Homo sapiens chromosome 1 clone RP5-833F5 map q22-23.3, ***
 SEQUENCING IN PROGRESS ***
 ACCESSION AL354988
 VERSION AL354988.4 GI:12539594
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 124096)
 Plumb, B.
 Direct Submision
 Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jan 25, 2001 this sequence version replaced gi:9797234.
 ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: d833f5
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 116353 bases at least Q40
 Consensus quality: 119600 bases at least Q30
 Consensus quality: 121395 bases at least Q20
 Insert size: 122596; sum-of-contigs
 Insert size: 161452; 28.1% error; agarose-IP
 Quality coverage: 3.30x in Q20 bases; sum-of-contigs quality
 coverage: 2.89x in Q20 bases; agarose-IP

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 3147 3246: contig of 3146 bp in length
* 3147 3246: gap of 100 bp
* 3247 6059: contig of 2813 bp in length
* 6060 6159: gap of 100 bp
* 6160 12590: contig of 6431 bp in length
* 12591 12690: gap of 100 bp
* 12691 33988: contig of 21298 bp in length
* 33989 34088: gap of 100 bp
* 34089 41478: contig of 7390 bp in length
* 41479 41578: gap of 100 bp
* 41579 63240: contig of 21662 bp in length
* 63241 63340: gap of 100 bp
* 63341 66406: contig of 3066 bp in length
* 66407 66506: gap of 100 bp
* 66507 74128: contig of 7622 bp in length
* 74129 74228: gap of 100 bp
* 74229 82511: contig of 8283 bp in length
* 82512 82611: gap of 100 bp
* 82612 87241: contig of 4630 bp in length
* 87242 87341: gap of 100 bp
* 87342 90226: contig of 2885 bp in length
* 90227 90326: gap of 100 bp
* 90327 96164: contig of 5838 bp in length
* 96165 96264: gap of 100 bp
* 96265 102960: contig of 6696 bp in length
* 102961 103060: gap of 100 bp
* 103061 105813: contig of 2753 bp in length
* 105814 105913: gap of 100 bp
* 105914 110611: contig of 4698 bp in length
* 110612 110711: gap of 100 bp
* 110712 124096: contig of 13385 bp in length.
  
```

FEATURES
 source
 1. 124096
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="q22-23.3"
 /clone="RP5-833F5"
 /clone_lib="RC1-5"
 1. 3146
 /note="assembly-fragment:00170
 fragment_chain:1"
 3247. 6059
 /note="assembly-fragment:00986
 fragment_chain:1"
 6160. 12590
 /note="assembly-fragment:00419
 fragment_chain:1"
 12691. 33988
 /note="assembly-fragment:01093

```

misc_feature 34089..41478
      /note="assembly-fragment:00516
      fragment_chain:1"
misc_feature 41579..63240
      /note="assembly-fragment:00023
      fragment_chain:2"
misc_feature 63341..66406
      /note="assembly-fragment:00261
      fragment_chain:2"
misc_feature 66507..74128
      /note="assembly-fragment:00198
      fragment_chain:3"
misc_feature 74229..82511
      /note="assembly-fragment:00734
      fragment_chain:3"
misc_feature 82612..87241
      /note="assembly-fragment:00437
      fragment_chain:4"
misc_feature 87342..90226
      /note="assembly-fragment:01058
      fragment_chain:4"
misc_feature 90327..96164
      /note="assembly-fragment:00438
      fragment_chain:5"
misc_feature 96265..102960
      /note="assembly-fragment:00639
      fragment_chain:5"
misc_feature 103061..105813
      /note="assembly-fragment:00045"
      fragment_chain:5"
misc_feature 105914..110611
      /note="assembly-fragment:00880"
      fragment_chain:5"
misc_feature 110712..124096
      /note="assembly-fragment:01070"
BASE COUNT 36796 a 22639 c 23902 g 39255 t 1504 others
ORIGIN
Query Match      83.9%; Score 187; DB 2; Length 124096;
Best Local Similarity 100.0%; Pred. No. 2.4e-46;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAATCACAGGAGATGATACAGCAATGGGCCATTAAAGTGTGTGTCATCTTGATT 60
      |||||
DB 105305 CTTACCTCTTGAGAGGGCCCTGAGTAAATTCATCTACCTGAACAACATGGCTAT 105364
QY 61 CTTACCTCTTGAGAGGGCCCTGAGTAAATTCATCTACCTGAACAACATGGCTAT 120
      |||||
DB 105355 CTTACCTCTTGAGAGGGCCCTGAGTAAATTCATCTACCTGAACAACATGGCTAT 105424
QY 121 GAAGCATTTGCTGTCATTCAGCCCAATGTGCCAGAGATGAACACTCATTCACAA 180
      |||||
DB 105425 GAAGCATTTGCTGTCATTCAGCCCAATGTGCCAGAGATGAACACTCATTCACAA 105484
QY 181 ATTAAGG 187
      |||||
DB 105485 ATTAAGG 105491

```

RESULT 6
 AL358950/c 164891 bp DNA HTG 23-JAN-2001
 LOCUS Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN
 DEFINITION PROGRESS ***
 ACCESSION AL358950
 VERSION AL358950.4 GI:12539689
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 164891)
 AUTHORS Plumb,B.


```

TITLE      Direct Submission
JOURNAL    Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire
           CE10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
           Requests: clonerequest@sanger.ac.uk
           On Jan 26, 2001 this sequence version replaced gi:9988471.
COMMENT    -----
           Genome Centre
           Center: Sanger Centre
           Center code: SC
           Web site: http://www.sanger.ac.uk
           Contact: humquerry@sanger.ac.uk
           Project Information
           Center project name: BA444C12
           Summary Statistics
           Assembly program: XGAP4; version 4.5
           Sequencing vector: plasmid; 108753; 100% of reads
           Chemistry: Dye-terminator Big Dye; 100% of reads
           Consensus quality: 161347 bases at least Q40
           Consensus quality: 162610 bases at least Q30
           Consensus quality: 163328 bases at least Q20
           Insert size: 163991; sum-of-ctrls
           Insert size: 163373; 6.6% error; agarose-fp
           Quality coverage: 5.19x in Q20 bases; sum-of-ctrls Quality
           coverage: 5.21x in Q20 bases; agarose-fp
           -----
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 10 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *
           * 1 34236: contig of 34236 bp in length
           * 34237 34336: gap of 100 bp
           * 34337 47372: contig of 13036 bp in length
           * 47373 47472: gap of 100 bp
           * 47473 67025: contig of 19553 bp in length
           * 67026 67125: gap of 100 bp
           * 67126 71239: contig of 4114 bp in length
           * 71240 71339: gap of 100 bp
           * 71340 83903: contig of 12564 bp in length
           * 83904 84003: gap of 100 bp
           * 84004 91357: contig of 7354 bp in length
           * 91358 91457: gap of 100 bp
           * 91458 96452: contig of 4995 bp in length
           * 96453 96552: gap of 100 bp
           * 96553 148189: contig of 51637 bp in length
           * 148190 148289: gap of 100 bp
           * 148290 159778: contig of 11489 bp in length
           * 159779 159878: gap of 100 bp
           * 159879 164891: contig of 5013 bp in length.
           *
           FEATURES
           source
           1. 164891
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /chromosome="1"
           /clone="RP11-444C12"
           /clone_11b="RPc1-11.2"
           1. 34236
           /note="assembly_fragment:00027
           fragment_chain:1
           clone_end:17
           vector_side:left"
           34337..47372
           /note="assembly_fragment:00312
           fragment_chain:1"
           47473..67025
           /note="assembly_fragment:01923
           fragment_chain:1"
           67126..71239
           /note="assembly_fragment:02094
           fragment_chain:1"
           71340..83903

```

misc_feature	/note="assembly:fragment:02213 fragment_chain:1" 84004..91357 /note="assembly:fragment:01292 fragment_chain:1" 91458..96452 /note="assembly:fragment:01391 fragment_chain:1" 96553..148189 /note="assembly:fragment:02045 fragment_chain:1" 148290..159778 /note="assembly:fragment:00427 fragment_chain:1" 159879..164891 /note="assembly:fragment:00773 fragment_chain:1 clone_end:SP6 vector_side:right"
BASE COUNT	48381 a 32179 c 32357 g 51067 t 907 others
ORIGIN	
Query Match	83.9%; Score 187; DB 2; Length 164891; Best Local Similarity 100.0%; Pred. No. 2,5e-46; Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GAATCACACAGGAGAGTGTACAGCATGTGGGCCATTTAACAGTCTGTCACTTTGANT 60
Db	80170 GAATCACACAGGAGAGTGTACAGCATGTGGGCCATTTAAGAGTCTGTCACTTTGANT 80111
OY	61 CTTACCTTTAGAGAGGGGCCCTAGTAATTCATCATTCACCTAGCTGAACACATGGCTAT 120
Db	80110 CTTACCTTTAGAGAGGGGCCCTAGTAATTCATCATTCACCTAGCTGAACACATGGCTAT 80051
OY	121 GAAGCATTTGCTGTGCATTCGACCCCAATGTGCCAGAGAAGATGAACACTATTACAA 180
Db	80050 GAAGCATTTGCTGTGCATTCGACCCCAATGTGCCAGAGAAGATGAACACTATTACAA 79991
OY	181 ATTAAGG 187
Db	79990 ATTAAGG 79984
RESULT 7	
AF039401	DNA PRI 14-DEC-1998
LOCUS	AF039401 35278 bp
DEFINITION	Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene, complete cds.
ACCESSION	AF039401
VERSION	AF039401.1
KEYWORDS	GI:4009459
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 35278) Gruber,A.D., Eldle,R.C., Ji,H.L., Schreuer,K.D., Puller,C.M. and Pauli,B.U.
TITLE	Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of Ca2+-activated Cl- channel proteins
JOURNAL	Genomics 54 (2), 200-214 (1998)
MEDLINE	99047526
REFERENCE	2 (bases 1 to 35278)
AUTHORS	Gruber,A.D., Eldle,R. and Pauli,B.U.
TITLE	Direct Submission
JOURNAL	Submitted (19-DEC-1997) Department of Pathology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES	Location/Qualifiers 1..35278
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606"


```

exon      /number=10
           /26644..26659
           /gene="hclcal"
exon      /number=11
           /27447..27708
           /gene="hclcal"
exon      /number=12
           /28766..28936
           /gene="hclcal"
exon      /number=13
           /31840..32079
           /gene="hclcal"
exon      /number=14
           /32919..33522
           /gene="hclcal"
exon      /number=15
           /11245 a 6870 c 7230 g 9933 t
BASE COUNT
ORIGIN
          *16
Query Match      83.1%; Score 185.4; DB 9; Length 35278;
Best Local Similarity 99.5%; Pred. No. 6.2e-46;
Matches 186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAATCAAGGAGAGATGTACAGCATGGGGCCATTTAAGAGTTCTGTTCATCTTGATT 60
    |||||
Db 2198 GGAATCAAGGAGAGATGTACAGCATGGGGCCATTTAAGAGTTCTGTTCATCTTGATT 2257

QY 61 CTTACCTCTCTAGAAGGGGCCCTGAGTAATTCATCTCATTCAGCTGACACAACTGCTAT 120
    |||||
Db 2258 CTTACCTCTCTAGAAGGGGCCCTGAGTAATTCATCTCATTCAGCTGACACAACTGCTAT 2317

QY 121 GAAGCATGTGTCGTGTGCATGACCCCAATGCGCCAGAAAGATGAACACTCATTCACACA 180
    |||||
Db 2318 GAAGCATGTGTCGTGTGCATGACCCCAATGCGCCAGAAAGATGAACACTCATTCACACA 2377

QY 181 ATAAAGG.187
    |||||
Db 2378 ATAAAGG 2384

RESULT 8
LOCUS AF095584 3079 bp mRNA MAM 13-OCT-2000
DEFINITION Sus scrofa epithelial chloride channel protein (AECC) mRNA,
complete cds.
ACCESSION AF095584
VERSION AF095584.1 GI:6002645
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE 1 (bases 1 to 3079)
AUTHORS Gaspar,K.J., Racette,K.J., Gordon,J.R., Loewen,M.E. and
Forsyth,G.W.
TITLE Cloning a chloride conductance mediator from the apical membrane of
porcine ileal enterocytes
JOURNAL Physiol. Genomics (Online) 3 (2), 101-111 (2000)
MEDLINE 20473747
PUBMED 11015605
REFERENCE 2 (bases 1 to 3079)
AUTHORS Gaspar,K.J., Gabriel,S.E., Racette,K.J. and Forsyth,G.W.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-1998) Veterinary Physiological Sciences,
University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4,
Canada
FEATURES
source 1..3079
location/Qualifiers
1..3079
/organism="Sus scrofa"
/db_xref="taxon:9823"
/tissue_type="ileal mucosa"
1..3079
gene

```


CDS

/gene="ABCC"
133..2886
/gene="ABCC"
/function="Involved in exocrine chloride conductance"
/product="epithelial chloride channel protein"
/protein_id="AAF0077.1"
/db_xref="GI:6002646"
/translation="MSFSSSLFVLVHLLEGASNSLIQLONGEGYIVADPNP
EDERLIONIKWTKASPYLPATEKERFEYKNVAILIPSMKAPPEYVPLETYNA
DYVTEPNPNPDGPTBOMGNGEKEGKILTPDPVAKKVLQVQPOGRVHEAH
LRMGVPEYNNQKRTLSNKKQPYICSAIKGTNYLPQCGGSCYTRKCRADRYTGL
FQKECEFIIDPOOSEKASIMFAOSIDTVEFEKYNHNKEAPDNQKCNLSRWTEVI
QDSEDFKRTPTQPPAPFSLQIGQRIYCLVLDKSGMTVGRKLKLNQKLESLA
LOTVEGAVGVAFDAAYVKSSELOVINSAPERDALSLPTAASGTSICSGLSRA
PLVKKRYPTDSEIVLITDGDNTISACEPYKONGATIIHTVALGSPAKLEELSL
MTGLOTVASDOAENNGLIDAFGLSSGNASORSIOLESQITLONNEMNNGYVY
DSTVGKDTFLITLERKLPSPFPGVPSGRSDSLVKNHKNMAYFOYPCIRAKVGM
KTSIDASSQTLTVSSRSATLPVPVYTSKMKDGTGFPSPMVYVYTHOCTPLIL
RAKVALIESENGKTVTLLELDNGAGADATKNDGYSRYETAYDANGRYSVKWMALG
VMTPRRAPPLMSGAMYIRGWTENGELIKWNPDPIDNKDLOGKQVCFRTASGGSFV
ASDVPRSPIDLPFPCIKTDLKAGIOGDNILMTWAPGDYDHRADRYTIRISNI
LIDRKENDSVQVNTTDLIPKRNSEVFEKPECTPPTGNDLFTAVQAVDKTNKS
EISNTAQSLEFLPEAPRPETPPETPPAPSLPCEPIOVNSTIIPGIIHKIMKMLGEIQL
SLA"

BASE COUNT 892 a 758 c 755 g 674 t
ORIGIN

Query Match 74.2%; Score 165.4; DB 4; Length 3079;
Best local Similarity 85.6%; Pred. No. 5.9e-40;
Matches 184; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 9 AGGAGATGTACAGCAATGGGCCATTAGAGTCTGTTCATCTTGATCTTCACCT 68
|||||
Db 117 AGGAGATGTACAGCAATGGGCCATTAGAGTCTGTTCATCTTGATCTTCACCT 176
QY 69 TCTAGAGGGGCCCTAGTAACTCATCTCAGCTGACGACAAACAGTGTATGAAGCAT 128
|||||
Db 177 TCTGAGAGGGGCTCAGTAACTCGCTCATCTCAGCTGACGACAAACAGTGTATGAAGCAT 236
QY 129 TGTCTGTTCATGACCCCATCTGCGAAGATGAAGCACTATCAACAATAAAGGA 188
|||||
Db 237 CGCATCTGCAATGACCCCATCTGCGAAGATGAAGCACTATCAACAATAAAGGA 296
QY 189 CATGTGACCCAGCACTCTGATCTGTTGAAG 223
|||||
Db 297 TATGTGACCAAGCATCTCATATCTGTTGAAG 331

RESULT 9
LOCUS AB017156 2937 bp mRNA ROD 10-NOV-1999
DEFINITION Mus musculus gob-5 mRNA, complete cds.
ACCESSION AB017156
VERSION AB017156.1 GI:3721911
KEYWORDS GOB-5.
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE 1 (sites)
AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.
TITLE Cloning and identification of the gene gob-5, which is expressed in
intestinal goblet cells in mice
JOURNAL Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)
MEDLINE 99100866
AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1998) to the DDBJ/EMBL/GenBank databases. Tohtu
Komiyama, Etsuo, JST, Genome Asymmetry Group, Doi Biosymmetry
Project, Tokodai 5-9-9, Tsukuba, Ibaragi 300-2635, Japan

(E-mail:tkom@bioa.erauto.lrc-net.co.jp, Tel:81-298-48-1515, Fax:81
298-47-8901)
FEATURES
source location/Qualifiers
1. 2937
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="goblet cell"
/dev_stage="adult"
/issue_type="Intestine"
15. 2756
/gene="gob-5"
15. 2756
/gene="gob-5"
/protein_id="BAA33743.1"
/db_xref="GI:3721912"

CDS

/translation="MSLSLSPFLVHLHLEGLVSLQLNNNGYIVADPNP
EDBALQIKNDWYQNSPYLPFATGKRFYFKNVAILIPSMKAPPEYVPLETYNA
DYVTSVSLGNDPEYTHEIGAGGEGIRIHLTPDLAKKLYOQPODRTVEHVAH
FRMGVPEYNNQKRTLSNKKQPYICSAIKGTNYLPQCGGSCYTRKCRADRYTGL
FQKECEFIIDPOOSEKASIMFNONINSVVECTEKNHNOEAPDNQKCNLSRWTEVI
QDSEDFKRTPTQPPAPFSLQIGQRIYCLVLDKSGMTVGRKLKLNQKLESLA
LOTVEGAVGVAFDAAYVKSSELOVINSAPERDALSLPTAASGTSICSGLSRA
PLVKKRYPTDSEIVLITDGDNTISCFDLVKGSAIHTVALGPPAAAELEQLS
KMTGGLQTVSSDOVONNGLVDAFAALISSGNAAIAQHSIOLESRGVNLQNNQNGSVI
VDSYVGRDTFLITWTHPTTIFMPDSEVONGFLIDTTKAYVLOVGTAKVGFMR
YSIOASSQTLTVTSRAASATLPITVTPVNNKNGKFPSPVYVYASIRBOGASPLR
ASVTALIESVNGKTVTLLELDNGAGADATKNDGYVSRYFTADANGRYSVKIRALAGV
TSORAPAPKRNKRAMTIDGWTIEDGEVRNMPREYSYVODKOLCESRTSSGSEVATN
VPAAPIDPLPFCQITDKASIOGDNILMTWAPGDYDHRASNYTIRISNTYD
LRHFTSVLOVNTTGLIPKRNSEVFEKPECTPPTGNDLFTAVQAVDKTNKSEI
SNARVSFVLPQEPPIPEDSPPCDINSITIPGIIHKIMKMLGEMOYTLGLH"
BASE COUNT 860 a 718 c 693 g 666 t
ORIGIN

Query Match 62.0%; Score 138.2; DB 10; Length 2937;
Best local Similarity 80.9%; Pred. No. 1.3e-31;
Matches 161; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 35 ATGGGGCATTTAAGAGTCTGTCTGATCTTCACTTCACTTCAAGAGGGCCCTG 84
|||||
Db 15 ATGGAATCTTGAAGAGTCTGTCTGATCTTCACTTCACTTCAAGAGGGCTCTG 74
QY 85 AGTAACTCACTCACTGCTGACGACAAACAGTGTATGAAGCAT 144
|||||
Db 75 AGTAACTCACTCACTGCTGACGACAAACAGTGTATGAAGCAT 134
QY 145 CCAATGTGCGAGAGTGAACACATCTCAACAATAAAGCATGTGACCGAGCA 204
|||||
Db 135 CAGGAGCTGCCGGAAGATGAAGCCCTCATCAACAATAAAGCATGTGACCGAGCC 194
QY 205 TCTCTGATCTGTTGAAG 223
|||||
Db 195 TCTCTGATCTGTTGAAG 213

RESULT 10
LOCUS AX192467 401 bp DNA PAT 15-AUG-2001
DEFINITION Sequence 34 from Patent WO0149716.
ACCESSION AX192467
VERSION AX192467.1 GI:15210431
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM human

REFERENCE 1 (bases 1 to 401)
AUTHORS Xu, J., Lodes, M. J., Secrist, H., Benson, D. R., Meagher, M. J.,
Stolk, J. A., King, G. E., Wang, T. and Jiang, Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 09:17:17 ; Search time 16681 Seconds
(Without alignments)
143.655 Million cell updates/sec

Title: US-09-049-696-1

Sequence: 1 GAAATCAGAGGAGATGTAC.....ATCTCTGATCTGTTTGAAG 223

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gp_est1:*
11: gp_est2:*
12: gp_hic:*
13: gp_gsa:*
14: em_gsa_fun:*
15: em_gsa_hum:*
16: em_gsa_inv:*
17: em_gsa_pln:*
18: em_gsa_pro:*
19: em_gsa_rtd:*
20: em_gsa_vrt:*
21: em_gsa_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	164.4	73.7	428	11	284017
2	144.4	64.8	1800	11	BF579071
3	138.2	62.0	745	11	BF579071
4	138.2	62.0	767	11	BF578246
5	138.2	62.0	806	11	BF578246
6	138.2	62.0	2915	12	AK007466
7	138.2	62.0	2933	12	AK008659
8	136.6	61.3	1069	11	BF324977
9	133.2	59.7	876	11	BF324977
10	125.6	56.3	998	11	BF583412
11	125.4	56.2	228	10	BF582485
12	123.8	55.5	231	10	BF581839

13	123	55.2	235	10	BB581817	BB581817
14	122.4	54.9	227	10	BB582380	BB582380
15	118.4	53.1	215	10	BB582408	BB582408
16	117.4	52.6	245	10	BB582806	BB582806
17	116.8	52.4	217	10	BB582407	BB582407
18	111	49.8	234	10	BB581749	BB581749
19	105.8	47.4	225	10	BB582566	BB582566
20	100.6	45.1	600	10	AL602665	AL602665
21	92.6	41.5	229	10	BB582543	BB582543
22	84.4	37.8	221	10	BB582559	BB582559
23	79.2	35.5	485	10	BB135550	BB135550
24	77.6	34.8	508	10	AT187977	AT187977
25	74.2	33.3	167	10	BB582528	BB582528
26	73.8	33.1	348	11	BF226118	BF226118
27	73.8	33.1	440	11	BI108239	BI108239
28	73.8	33.1	509	10	AM323607	AM323607
29	73.8	33.1	568	10	BB375531	BB375531
30	73.8	33.1	607	11	BB175323	BB175323
31	73.8	33.1	674	10	BE380998	BE380998
32	73.8	33.1	761	11	BI251115	BI251115
33	73.8	33.1	827	11	BB176202	BB176202
34	73.8	33.1	912	11	BB913371	BB913371
35	73.8	33.1	947	11	BB244665	BB244665
36	73.8	33.1	961	11	BB173925	BB173925
37	73.8	33.1	1008	11	BB247592	BB247592
38	73.4	32.9	275	10	BB572124	BB572124
39	73	32.7	157	10	BB583300	BB583300
40	72.2	32.4	519	11	BB798485	BB798485
41	72.2	32.4	638	10	AT190317	AT190317
42	72.2	32.4	759	11	BB248072	BB248072
43	72.2	32.4	1153	11	BF101091	BF101091
44	70.6	31.7	884	11	BF102315	BF102315
45	69.6	31.2	147	10	BB581674	BB581674

ALIGNMENTS

RESULT 1
LOCUS 284017
DEFINITION SS284017 Porcine small intestine cDNA library Sus scrofa clone
c12a03 5', mRNA sequence.
ACCESSION 284017
VERSION 284017.1 GI:1806336
KEYWORDS EST.
SOURCE Sus scrofa
ORGANISM Sus scrofa
Pig.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE 1 (bases 1 to 428)
Wintzer, A.K., Fredholm, M. and Davies, W.
TITLE Evaluation and characterization of a porcine small intestine CDNA
library: analysis of 839 clones
JOURNAL Mamm. Genome 7 (7), 509-517 (1996)
MEDLINE 96327607
COMMENT Contact: A.K. Wintzer
Department of Animal Science and Animal Health, Division of Animal
Genetics, The Royal Veterinary and Agricultural University
Bulowsvej 13, 1870 Frederiksberg C, Denmark.
FEATURES
SOURCE 1. 428
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="c12a03"
/clone_lib="Porcine small intestine cDNA library"
/note="directionally cloned cDNA in X1-blue MRF."
BASE COUNT 134 a 87 c 100 g 102 t 5 others
ORIGIN
Query Match 73.7%; Score 164.4; DB 11; Length 428;
Best Local Similarity 83.8%; Pred. No. 5e-39;

Matches 186; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

OY 2 AATCAAGGAGATGTACAGCAATGGGGCCATTAAAGATTCTGTTCATCTTGAATTC 61
DB 3 AACCCAGAGGAGATGCACAGCCATGAGGATTCATTAGAGTTCCGCTTCATCTGCTGC 62
OY 62 TTCACCTTCTAGAGGGGCGCTGAGTAATTCATCTATTCAGTGAACAAACATGCGTATG 121
DB 63 TTCACCTTCTAGAGGGGCGCTGAGTAATTCATCTATTCAGTGAACAAACATGCGTATG 122
OY 122 AAGGCATTCCTGTTCAGATTCAGCCCAATGTGCGAGAAAGTGAACATCATCAACAA 181
DB 123 AAGGCATTCCTGTTCAGATTCAGCCCAATGTGCGAGAAAGTGAACATCATCAACAA 182
OY 182 TAAAGGACATGTGTACCCAGGACATCTCTGTATCTGTTTGAAG 223
DB 183 TAAAGGATATGTGTACCCAGGACATCTCTGTATCTGTTTGAAG 224

RESULT 2

LOCUS BF579071 1800 bp mRNA EST 12-DEC-2000
DEFINITION 602096001F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4215841 5',
mRNA sequence.
ACCESSION BF579071
VERSION BF579071.1 GI:11652783
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1800)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-ri@mail.nih.gov
Tissue procurement: Jeffrey E. Green, M.D.
cDNA library preparation: Life Technologies, Inc.
cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM9792 row: b column: 02
High quality sequence stop: 559.
Location/Qualifiers
1. 1800
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="IMAGE:4215841"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source

BASE COUNT 631 a 472 c 492 g 205 t
ORIGIN

Query Match 64.8%; Score 144.4; DB 11; Length 1800;
Best local similarity 78.9%; Pred. No. 6e-33;
Matches 172; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

OY 6 CACAGGAGATGTACAGCAATGGGGCCATTAAAGATTCTGTTCATCTTGAATTC 65
DB 16 CAGAGGAAAGTGCAGAGTGAAGTCTTTGAAGACTCTGCTTCCTTGAATCTCCA 75
OY 66 CTTCTAGAGGGGCGCTGAGTAATTCATCTATTCAGTGAACAAACATGCTATGAGG 125
DB 76 CTTCTAGAGGGGCGCTGAGTAATTCATCTATTCAGTGAACAAACATGCTATGAGG 135

OY 126 CATTCGCTGTTCAGATTCAGCCCAATGTGCGAGAAAGTGAACATCATCAACAA 185
DB 136 CATTCGCTGTTCAGATTCAGCCCAATGTGCGAGAAAGTGAACATCATCAACAA 195
OY 186 GGACATGTGTACCCAGGACATCTCTGTATCTGTTTGAAG 223
DB 196 GGACATGTGTACCCAGGACATCTCTGTATCTGTTTGAAG 233

RESULT 3

LOCUS BG963539 745 bp mRNA EST 12-JUN-2001
DEFINITION 602831217F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4985741 5',
mRNA sequence.
ACCESSION BG963539
VERSION BG963539.1 GI:14351176
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 745)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-ri@mail.nih.gov
Tissue procurement: Jeffrey E. Green, M.D.
cDNA library preparation: Life Technologies, Inc.
cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM10994 row: a column: 06
High quality sequence stop: 745.
Location/Qualifiers
1. 745
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="IMAGE:4985741"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source

BASE COUNT 236 a 173 c 172 g 164 t
ORIGIN

Query Match 62.0%; Score 138.2; DB 11; Length 745;
Best local similarity 80.9%; Pred. No. 4e-31;
Matches 161; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

OY 35 ATGGGCGCTTAAAGATTCGTTCATCTGATTCCTACCTTCAAGAGGGGCGCTG 84
DB 15 ATGGATCTTTAAAGAGTCTCTCTCTGATCTCTCCACCTCTGGAAGATTTCTG 74
OY 35 ATGATTCATCTATTCAGTGAACAAATGCTATGAGGAGCTGCTGCAATGCAG 144
DB 75 ATGATTCATCTATTCAGTGAACAAATGCTATGAGGAGCTGCTGCAATGCAG 134
OY 135 CACGAGGTCGCGAGAGATGAACCTTCATCAACATTAAGGACATGCTGAGGCC 204
DB 135 CACGAGGTCGCGAGAGATGAACCTTCATCAACATTAAGGACATGCTGAGGCC 194
OY 205 TCTCTATCTGTTTGAAG 223
DB 135 TCTCTATCTGTTTGAAG 213

RESULT	4	767 bp	mRNA	EST	12-DEC-2000
LOCUS	BF578246				
DEFINITION	602093073p1	NCI_CGAP_Co24	Mus musculus	CDNA clone	IMAGE:4207386 5'
ACCESSION	BF578246				
VERSION	BF578246.1	GI:11651958			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 767)				
TITLE	Nih-MGC http://mgc.nci.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM9770 row: a column: 19 High quality sequence stop: 679. Location/Qualifiers 1..767 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone_image:4207386 /clone_11b="NCI_CGAP_Co24" /lab_host="DH10B (TI phage-resistant)" /note="Organ: colon; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.6 kb. Constructed by life Technologies. Note: this is a NCI_CGAP Library." BASE COUNT 235 a 176 c 180 g 176 t ORIGIN				
FEATURES	source				
Query Match	62.0%;	Score 138.2;	DB 11;	Length 767;	
Best Local Similarity	80.9%;	Pred. No. 4e-31;	38;	Indels 0;	Gaps 0
Matches 161:	Conservative	0;	Mismatches	38;	Indels 0;
OY	25	ATGGGCAATTAAGATCTGTGTTCACTTGATCTTCACTTGTAGAAAGGAGCCG 84			
DB	23	ATGGAAATCTTTAAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 82			
OY	85	AGTAATTCACATTCATGCTGAGTAACAACAATGGCTATGAAGGCAATGTGTCATGCAT 144			
DB	83	AGTGATGCTCCCTATCCCACTGAACAACAACGCGCTATGAGGCGCATGCTCATCCCATAGAC 142			
OY	145	CCCAATGGCCGGAAGATGAACACATTCATTAACAATTAAGACATGGTGAACCGGCA 204			
DB	143	CACGACGTGCCGGAAGATGAAGCCCTCATTTCAACACATTAAGACATGGTGAACGCC 202			
OY	205	TCTCTATCTCTTTGAAG 223			
DB	203	TCTCATACCTGTTGAAG 221			
RESULT	5				
LOCUS	BG962080	806 bp	mRNA	EST	12-JUN-2001
DEFINITION	602826824p1	NCI_CGAP_Co24	Mus musculus	CDNA clone	IMAGE:4981853 5'
ACCESSION	BG962080				
VERSION	BG962080.1	GI:14349717			
KEYWORDS	EST.				

SOURCE		house mouse.
ORGANISM		Mus musculus
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 806)
AUTHOR(S)		NIH-MGC http://mgc.nci.nih.gov/.
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL
COMMENT		Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LHAM10983 row: o column: 06 High quality sequence stop: 713. Location/Qualifiers 1..806 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:4981853" /clone_id="NCI_CGAP_C024" /lab_host="DH10B (TI phage-resistant)" /note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI-CGAP library."
BASE COUNT	248 a	194 c 183 g 181 t
ORIGIN		
Query Match	62.0%; Score 138.2; DB 11; Length 806;	
Best Local Similarity	80.9%; Pred. NO. 4e-31;	
Matches: 161; Conservative	0; Mismatches 38; Indels 0; Gaps	0;
CY	45	ATGGGCGCATTAAAGAGTTCTGTTCATTCATTCACCTCTGAAGGGCCCTG 84
D3	47	ATGGAGCTTTGAAGAGTGCTCTCTCTCTCTTGATGCCACCTCTGGAAGGATTTCTG 76
CY	35	AGTAATTCATCATTCACCTGACAACAAATGGCTATGTAAGGCATTTCTGTCATTCAGC 144
D3	77	AGTAGTCTCCCTATCCACTGACAAACAACGCGTATGAGGCGATGTCATGCCCAATAAC 136
CY	145	CCCAGTGGCCAGAAGATGAACACTCTTCACAAATAAAGACATGTGACCAGCA 204
D3	117	CACGACGTGGCCGGAAGATGAACCTCTCATTCACACATTAAGACATGATGACTAGGCC 196
CY	325	TCTGTATCTGTGTTGAAG 223
D3	137	TCTCATACCTGTTGAAG 215
RESULT	6	
Locus	AK007466	mRNA HTC 05-JUN-2001
DEFINITION		Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810012P03, full insert sequence.
ACCESSION	AK007466.1	GI:12841032
VERSION:		
KEYWORDS:		Cap. trapper.
SOURCE		Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library clone:1810012P03.
ORGANISM		Mus musculus
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2915)
AUTHORS		Carninci, P. and Hayashizaki, Y.
TITLE		High-efficiency full-length cDNA cloning

OY	25	ATGGGCGCATTTAAGATTGTGTGTATCTTGATTTCTTACCTTAGAAGGGGCCTG	84
Dd	27	ATGGAAATTTGAAGAAGTCCTTTCTTCTCTTGATFCCITCACCCTTCGGAAGGAGTTCTG	86
OY	85	AGTAATTCATCAATTCAGCTGACGATAACAACATGGCTATGAGGCAATGTTCGTTGCATTCGAC	144
Dd	87	AGTAGTCCCTCATTCACACTGACACACACAGGCTATGAGGCGATCGATCGCCATGAC	146
OY	145	CCCAGATGTGCCAAGAAAGAACACTCATTCACAAATTAAGAGACATGGTAGCCAGSCA	204
Dd	147	CACGACGGCGCCGGAAGATGAAGCCCTCATTTACACACATTAAGGACATGTGATCTACGACC	206
OY	205	TCTCGTATCTGTTTGAAG	223
Dd	207	TCTCATCTACCTGTTTGAAG	225
RESULT	9		
BF234977			
LOCUS	DEFINITION	BF234977 876 bp mRNA EST 14-NOV-2000 602027529P1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4162662 5'	
ACCESSION	VERSION	BF234977.1 GI:11147045	
KEYWORDS	SOURCE	EST.	
ORGANISM	SOURCE	house mouse. Mus musculus	
REFERENCE	AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
JOURNAL	TITLE	Nih-Mgc http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT	JOURNAL	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LLM9445 row: j column: 07 High quality sequence stop: 672.	
FEATURES	source	Location/Qualifiers 1..876 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:4162662" /clone_lib="NCI_CGAP_L19" /lab_host="DH10B (TI phage-resistant)" /note="Organ: liver; Vector: pcwv-sPORT6; site:1; NotI; site:2; SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: This is a NCI_CGAP Library."	
BASE COUNT	ORIGIN	269 a 223 c 204 g 180 t	
Query Match		59.7%; Score 133.2; DB 11; Length 876;	
Best Local Similarity		80.4%; Pred. No. 1.3e-29;	
Matches 156; Conservative		0; Mismatches 38; Indels 0; Gaps 0;	
OY	25	ATGGGCGCATTTAAGATTGTGTGTATCTTGATTTCTTACCTTAGAAGGGGCCTG	84
Dd	14	ATGGAAATCTTTGAAGAGTCCTGTCCTCTTATCTTCACCTTTGGAAGAGTTCTG	73
OY	85	AGTAATTCATCAATTCAGCTGACGATAACAACAGGCTATGAGGCAATGTTCGTTGCATTCGAC	144
Dd	74	AGTAGATCCCTCATTCACACTGACACACAGGCTATGAGGCGATCGATCCCATGAC	133
OY	145	CCCAGATGTGCCAAGAAAGAACACTCATTCACAAATTAAGAGACATGTGATCCAGCCAGCA	204

D0	194	CACGACGTGGCGGAAGATGAAAGCCCTCATTCACACATTAAGAAGACATGTACTAGGCC	193
Q7	205	TCTCCTGATCTGTT	218
D0	194	TCTCCTGATCTGTT	207
RESULT	10		
	EF583412		
LOCUS	602097421F1	NCL_CGAP_CO24	Mus musculus cDNA clone IMAGE:4217548 5'
DEFINITION	mRNA; sequence.		
ACCESSION	BF583412		
VERSION	BF583412.1		
KEYWORDS	EST		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.		
TITLE	(bases 1 to 998)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: rgs@phs-remail.nih.gov		
	Tissue/Procurement: Jeffrey E. Green, M.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.lnl.gov		
	Plate: LLM9796 row: 1 column: 05		
	High quality sequence stop: 658.		
FEATURES	Location/Qualifiers		
SOURCE	1..998		
	/organism="Mus musculus"		
	/strain="FWB/N"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:4217548"		
	/clone.lib="NCL_CGAP_CO24"		
	/lab_host="DH10B (TI phage-resistant)"		
	/note="Origin: colon; Vector: pCMV-SPORT6; Site_1: NCI;		
	Site_2: Salt; Cloned unidirectionally. Primer: Oligo dt.		
	Average insert size 1.6 kb. Constructed by Life		
	Technologies. Note: this is a NCL_CGAP library."		
BASE COUNT	336 a 238 c 250 g 194 t		
ORIGIN			
	Query Match	56.3%; Score 125.6; DB 11; Length 998;	
	Best Local Similarity	80.0%; Pred. No. 2.5e-27;	
	Matches 160; Conservative	0; Mismatches 39; Indels 1; Gaps 1	
Q7	25	ATGGGGCCATTAAAGTCTGTGTCATCACTTGATCTTCACCTTAGAAGGGGCCCTG	84
D0	21	ATGCAATCTTTTGAAGAGTCTGTCTTCCTTGATCTTCACCTTCGAGGAGGATTTCTG	70
Q7	35	AGTAATTCACCTCACTTAGCTGAGACAACAATGGCTATGAGAAGGCACTTGTCGTCATCGAC	144
D0	71	AGTAGATCCCTATATCCAATGACACACACACGCGTTATGAGGGCATCGTCATCGCCATAGAC	130
Q7	115	CCCAATGTGGCCAGACAGATGAAACAGCTCATTCAA-CAAATAAGGACATGCTGACCAGGC	203
D0	131	CACGACGTGGCGGAAGATGAAAGCCCTCATTCACACATTAAGAAGACATGTACTAGGCC	190
Q7	214	ATCTCTGATCTGTTGAAG	223
D0	131	CTCTCCAGACCTGTTGAAG	210
RESULT	11		

Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

Source

Location/Qualifiers
 1..215
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="9030403K21"
 /clone_lib="RIKEN full-length enriched, adult male colon"
 /sex="male"
 /tissue_type="colon"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATTCGATCAAGAGCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGATCAAGAGCTTTTCTTNN 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified phagescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

BASE COUNT 63 a 56 c 47 g 49 t
 ORIGIN

Query Match

Best Local Similarity 53.1%; Score 118.4; DB 10; Length 215;
 Matches 143; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```

OY 25 ATGGGGCATTTAAGATTCTGTTCATCTTATCTTCACCTTCAGAGGGGCCCTG 84
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 29 ATGGCAATCTTTGAAATCTCTGCTTCCTTATCTTCACCTTCAGAGAAATCTTG 88
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 85 AGTAATTCACCTCATTCAGCTGAACAACAATGGCTATGAAGCATGTGCTGCAATCGAC 144
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 89 AATGAGTCCCTCAVCCAACTGACAAACGAGGTATGAGGCGATCGCATCGCATAGAC 148
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 145 CCCAATGTGCCAAGATGAACACATCATTCACCAATAAAGACATCGTGAACCCAGCA 204
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 149 CACGATGTGCCGAAATGAAGCCCTCATTCACACATTAAGACATGTGACTCAGGCC 208
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 205 TCTC 208
    |||||
DB 209 TCTC 212
    |||||

```

Search completed: April 3, 2002, 20:27:06
 Job time: 40189 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:45 ; Search time 16681 Seconds

(without alignments)
155.894 Million cell updates/sec

Title: us-09-049-696-16

Perfect score: 242
Sequence: 1 GTTTATTCCTCCACAGACTC.....ATGTTATTTAGACTTCTCTG 242

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estlom:*
5: em_estpl:*
6: em_estbda:*
7: em_estbro:*
8: em_estov:*
9: em_hlc:*
10: gb_est1:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_tod:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
C 1	242	100.0	450	10	AM009764	ws87b05.x	AM009764	ws87b05.x
C 2	242	100.0	450	10	AA508854	n122h10.s	AA508854	n122h10.s
C 3	242	100.0	712	10	AM009763	ws87b04.x	AM009763	ws87b04.x
C 4	242	100.0	716	10	A1660234	ws68g02.x	A1660234	ws68g02.x
C 5	226.8	93.7	392	10	A1687981	wa77f06.x	A1687981	wa77f06.x
C 6	213	86.0	466	11	BG195179	RS114358	BG195179	RS114358
C 7	211.4	87.4	561	10	A1802756	wf18a04.x	A1802756	wf18a04.x
C 8	201.2	83.1	255	11	D25727	HMG504094	D25727	HMG504094
C 9	199	82.2	504	10	A1802693	wf17c01.x	A1802693	wf17c01.x
C 10	199	82.2	734	10	A1660957	wf20d08.x	A1660957	wf20d08.x
C 11	194.6	80.4	501	10	A1721275	as88h08.x	A1721275	as88h08.x
C 12	183.8	76.0	335	10	A1721121	as73d08.x	A1721121	as73d08.x

13	181	74.8	378	10	AA296955	EST112726	AA296955	EST112726
14	155.8	64.4	161	10	A1582072	ar96a11.x	A1582072	ar96a11.x
15	92	38.0	388	10	AA692521	vt59g08.r	AA692521	vt59g08.r
16	92	38.0	388	10	AA734161	vs19g06.r	AA734161	vs19g06.r
17	92	38.0	418	10	AA596289	vs26a08.r	AA596289	vs26a08.r
18	92	38.0	421	10	AA691586	vs12g06.r	AA691586	vs12g06.r
19	92	38.0	431	10	AA238284	my34g09.r	AA238284	my34g09.r
20	92	38.0	463	10	AA839323	vc04f07.r	AA839323	vc04f07.r
21	92	38.0	482	10	AA711228	vt75b06.r	AA711228	vt75b06.r
22	92	38.0	488	10	AA733415	vt75b06.r	AA733415	vt75b06.r
23	92	38.0	2915	12	AK007466	Mus muscu	AK007466	Mus muscu
24	92	38.0	2933	12	AK008659	Mus muscu	AK008659	Mus muscu
25	91.6	308	308	10	AV049139	AV049139	AV049139	AV049139
26	90.4	295	10	AV050167	AV050167	AV050167	AV050167	AV050167
27	89.2	36.9	812	11	BG568878	602836186	BG568878	602836186
28	87	36.0	552	10	AA688953	vs04h07.r	AA688953	vs04h07.r
29	84	34.7	732	11	BG693719	602831354	BG693719	602831354
30	83.8	306	10	AV059168	AV059168	AV059168	AV059168	AV059168
31	82.6	34.1	270	10	BB072582	BB072582	BB072582	BB072582
32	81.6	33.7	278	10	AV058023	AV058023	AV058023	AV058023
33	81.4	33.6	331	10	AV051601	AV051601	AV051601	AV051601
34	81	33.5	268	10	AV375939	AV375939	AV375939	AV375939
35	81	33.5	280	10	AV067322	AV067322	AV067322	AV067322
36	80	33.1	283	10	AV076160	AV076160	AV076160	AV076160
37	79.6	32.9	268	10	AV373378	AV373378	AV373378	AV373378
38	78.4	32.4	254	10	AV375072	AV375072	AV375072	AV375072
39	77.6	32.1	256	10	AV378819	AV378819	AV378819	AV378819
40	77.4	32.0	237	10	AV369383	AV369383	AV369383	AV369383
41	77.4	32.0	259	10	AV373600	AV373600	AV373600	AV373600
42	76.4	31.6	270	10	AV373097	AV373097	AV373097	AV373097
43	76.2	31.5	248	10	AV368948	AV368948	AV368948	AV368948
44	76.2	31.5	268	10	AV373431	AV373431	AV373431	AV373431
45	76	31.4	258	10	AV079760	AV079760	AV079760	AV079760

ALIGNMENTS

RESULT 1
LOCUS AM009764/C
DEFINITION ws87b05.x1 NCI_CGAP_C03 Homo sapiens cDNA clone IMAGE:2504913 3'
ACCESSION AM009764
VERSION AM009764.1 GI:5858542
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 450)
NCT-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. consortium/ILN at:
www.bio.lnlnl.gov/bbrp/image/image.html
Insert Length: 387 Std Error: 0.00
Seq primer: -40bp from Gibco.
FEATURES
source
1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2504913"
/clone_lib="NCI_CGAP_C03"


```

/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from 12 pooled bulk tumor samples and primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library went through one round of
normalization."

```

Query Match	100.0%	Score 242	DB 10	Length 450
Best Local Similarity	100.0%	Pred. No.	2e-45	
Matches 242	Conservative	0	Mismatches	0
			Incls	0
			Gaps	0

Oy	1	GTATTTCCTCCACAGACTCCGCCAGAGACACCTAGTCTCGATGAAACGTCCTCTTG	60
Db	373	GTTTATTCTCCACAGACTCCGCCAGAGACCTAGTCTCGATGAAACGTCCTCTTG	314
Oy	61	TCTCATATTTTATATCAACACACCACCATTTCCGCGATTCACATTTTAAATTTATNGGAA	120
Db	313	TCTCATATTTTATATCAACACACCACCATTTCCGCGATTCACATTTTAAATTTATNGGAA	254
Oy	121	GTGGATAGGAGAACTGCAGCTGTCAATAGCCTAGCGCTGATTTTGTCTGATTAATATAA	180
Db	253	GTGGATAGGAGAACTGCAGCTGTCAATAGCCTAGCGCTGATTTTGTCTGATTAATATAA	194
Oy	181	ATAAATCATTCATCTTTTTCGATTAATAAATTTCTTAAATNGTATTTTACACTTCCT	240
Db	193	ATAAATCATTCATCTTTTTCGATTAATAAATTTCTTAAATNGTATTTTACACTTCCT	134
Oy	241	GT 242	
Db	133	GT 132	

RESULT	2
AA508854/c	
LOCUS	
DEFINITION	
AA508854	490 bp mRNA
n12nh10.s1 NCI_CGAP_Co4 Homo sapiens cDNA clone IMAGE:568803	EST 18-AUG-1997
mRNA sequence.	3'

ACCESSION	AA508854	
VERSION	AA508854.1	GI:2246357
KEYWORDS	EST.	

ORGANISM:

REFERENCE	(bases 1 to 490)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strusberg, Ph.D.

Ph.D., Michael R. Emmert-Buck, M.D.
Tissue Procurement: Elias Campo, M.D.,
Email: ccampo-t@mail.nih.gov
Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CCAP clone distribution information can be
found through the I M J C B Consortium/INT at:

Round through the I.M.A.U.E. consolidation at
www-bio.lnl.gov/bbrp/image/image.html
root 1000th 956 Std Error: 0.00

Insert length: 858 Seq Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 404.

Location/Qualifiers

```
source 1. .490
/organism="Homo sapiens"
```

```

BASE COUNT      155 a
ORIGIN
/db_xref="taxon:9606"
/clone="IMAGE:968803"
/clone_1bp="Not_CGAP_CO4"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/notes="Vector: pRTT3-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from pooled color
tumor tissue, and was then primed with a Not I - oligo(dT)
primer. Double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pRTT3 vector.
This library is not normalized. Library constructed by
Bento Soares and M. Fatima Ronaldo."
155 a      82 c      89 g      164 t

```

```

Query Match 100.0% Score 242 DB 10 Length 450;
Best Local Similarity 100.0% Pred. No. 1.9e+45
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTTTTTCCTCCACAGACTCGGCAGAGACACTAGTCTGATGAAACGTCTGCTCTTG 60
DB 478 GTTTATTTCTCCACAGACTCGGCAGAGACACTAGTCTGATGAAACGTCTCTCTTG 349
QY 51 TCTTAATTTTCATATCAACAGACACCATTCTCTGGCATTCACATTTTAAATAATATATGAGGA 120
DB 348 TCTCAATATTTATATTCACAACAGACACCATTCTCTGGCATTCACATTTTAAATAATATGAGGA 289
QY 121 GTGATATGAGAGACTGCAGCTGTCAATATAGCCTAGCGCTGATTTTGTCCAGATAAATAA 180
DB 298 GTGGATATGAGAGACTGCAGCTGTCAATATAGCCTAGCGCTGATTTTGTCCAGATAAATAA 229
QY 181 ATAAATCATATCATCTCTTTTGTGATTTATAAATTTTCTAAATATGATTTTAAGACTTCC 240
DB 228 ATAAATCATATCATCTCTTTTGTGATTTATAAATTTTCTAAATATGATTTTAAGACTTCC 169
QY 241 GT 242
DB 168 GT 167

RESULT	3				
AM009763/c					
LOCUS					
DEFINITION					
AM009763	712 bp	mRNA	EST	08-MAR-2000	
ws87bd4.x1	NCI-GAD-Co3	Homo. sapiens	cDNA clone IMAGE:2504911		
similar to	TR:088626	088626	GOB-5 PROTEIN. ;	mRNA sequence.	

ACCESSION	AW009763
VERSION	AW009763.1
KEYWORDS	EST.
	GI:5858541

ORGANISM

REFERENCE
AUTHORS
TITLE:
JOURNAL
COMMENT:

Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Canis; Canis; Homo.
1 (bases 1 to 712)
NCI-CCSP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

Email: CGAPs-Email.hn.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arranging: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 807 std Error:
Gen primer: -400bp from Gibco

seq primer: -400 from 5' end
High quality sequence stop: 451.

FEATURES
source

Location/Qualifiers
1. 392
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2302211"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NDH19M, testis NHT, and B-cell NCI-GAP-GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
125 a 69 c 67 g 131 t

ORIGIN

Query Match 93.7%; Score 226.8; DB 10; Length 392;
Best Local Similarity 99.1%; Pred. No. 5.8e-42;
Matches 228; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTTATTCCTCCACAGACTCCGCCAGAGACACCTAGTCTGATGAACGTCCTCTTG 60
|||||
Db 372 GTTATTCCTCCACAGACTCCGCCAGAGACACCTAGTCTGATGAACGTCCTCTTG 313
|||||
QY 61 TCTATATTCATATCAACAGCAGCCATCTGCTGATTCATTTAAATATATGGA 120
|||||
Db 312 TCTATATTCATATCAACAGCAGCCATCTGCTGATTCATTTAAATATATGGA 253
|||||
QY 121 GTGATAGGAGACAGCTGCTGATTAAGCTAGGGCTGAATTTTGTGAGATTAATA 180
|||||
Db 252 GTGATAGGAGACAGCTGCTGATTAAGCTAGGGCTGAATTTTGTGAGATTAATA 193
|||||
QY 181 ATAAATCATTCATCCTTTTGTGATTAATAATTTCTTAAATGATATT 230
|||||
Db 192 ATAAATCATTCATCCTTTTGTGATTAATAATTTCTTAAATGATATT 143
|||||

RESULT 6
Bg195179 466 bp mRNA EST 21-APR-2001
LOCUS Bg195179
DEFINITION RST14358 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION Bg195179
VERSION Bg195179.1 GI:13716866
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 466)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Dahl,T., Thornton,M., Ramchandran,R., Whittington,J.,
Lerner,L., Krashoc,D., McElligott,K., Clark,S., Mays,R., Smith,E.,
Veloso,N., Hess,J., Cothran,K., Lo,K., Offenbacher,J., Danzig,J.
and Ducar,M.
TITLE Creation of Genome-wide Protein Expression Libraries using Random
Activation of Gene Expression
JOURNAL Nat. Biotechnol. 19 (5), 440 (2001) in press
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scai@atersys.com

FEATURES
source High quality sequence stop: 466.
Location/Qualifiers
1. 466

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT
155 a 92 c 74 g 144 t

ORIGIN

Query Match 88.0%; Score 213; DB 11; Length 466;
Best Local Similarity 100.0%; Pred. No. 7.7e-39;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTATTCCTCCACAGACTCCGCCAGAGACACCTAGTCTGATGAACGTCCTCTTG 60
|||||
Db 245 GTTATTCCTCCACAGACTCCGCCAGAGACACCTAGTCTGATGAACGTCCTCTTG 304
|||||
QY 61 TCTATATTCATATCAACAGCAGCCATCTGCTGATTCATTTAAATATATGGA 120
|||||
Db 305 TCTATATTCATATCAACAGCAGCCATCTGCTGATTCATTTAAATATATGGA 364
|||||
QY 121 GTGATAGGAGACAGCTGCTGATTAAGCTAGGGCTGAATTTTGTGAGATTAATA 180
|||||
Db 365 GTGATAGGAGACAGCTGCTGATTAAGCTAGGGCTGAATTTTGTGAGATTAATA 424
|||||
QY 181 ATAAATCATTCATCCTTTTGTGATTAATA 213
|||||
Db 425 ATAAATCATTCATCCTTTTGTGATTAATA 457
|||||

RESULT 7
A1802756/c 561 bp mRNA EST 18-DEC-1999
LOCUS A1802756
DEFINITION w18a04.x1 Soares Dieckgrafe.colon_NHUC Homo sapiens cDNA clone
IMAGE:2350926 3' similar to TR:088826 O88826 G08-5 PROTEIN. // mRNA
sequence.
ACCESSION A1802756
VERSION A1802756.1 GI:5368228
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 561)
NCI-GAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT Email: cga@bbs-rtmail.nih.gov
Contract: Robert Strausberg, Ph.D.
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 840 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 398.
Location/Qualifiers
1. 561

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2350926"
/clone_lib="Soares-Dieckgrafe.colon_NHUC"
/tissue_type="colonic mucosa from 5 ulcerative colitis patients"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTACCAATCTGAAGTGGAGCGCGCTAGCTTTTGTGATTTT 3'),

QY	61	TCCTAATATTCATATTCACAGACAGACCAATTCCTGGCATTCACATTTTAAATTAATGTGGAA	120
Db	151	TCCTAATATTCATATTCACAGACAGACCAATTCCTGGCATTCACATTTTAAATTAATGTGGAA	92
QY	121	GTGGATAGAGAACTGCACCTGTCAATAGCCGAGGCGTGAATTTTGTTCAGATTAATAAA	180
Db	91	GTGGATAGAGAACTGCACCTGTCAATAGCCGAGGCGTGAATTTTGTTCAGATTAATAAA	32
QY	181	ATAAATCATTCATCCCTTTT	199
Db	31	ATTAATCATTCATCCCTTTT	13
RESULT	10		
LOCUS	A1660957	734 bp	EST
DEFINITION	wf20d08.x1 Soares_Dieckgraefe.colon_NHUC Homo sapiens cDNA clone IMAGE:2351151 3' similar to TR:088826 088826 GDB-5 PROTEIN. ; , mRNA sequence.		
ACCESSION	A1660957		
VERSION	A1660957.1	GI:4764540	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 734)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997)		
FEATURES	Contact: Robert Strausberg, Ph.D. Email: c9gpb5-remail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 820 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 477. Location/Qualifiers 1..734 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2351151" /clone_lib="Soares_Dieckgraefe.colon_NHUC" /tissue_type="colonic mucosa from 5 ulcerative colitis patients" /lab_host="DH10B (phage-resistant)" /note="Organ: colon; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAAGTGGAGCGCGCCGTACCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieck@wustl.edu); colonic mucosa represents a range of disease involvement from mild cryptitis to severe ulceration, fibrosis, and degeneration. Library constructed by Bento Soares and M. Ratlma Bonaldo.		
BASE COUNT	200 a	143 c	159 g
ORIGIN		228 t	4 others
Query Match	82.2%	Score 199;	DB 10;
Best Local Similarity	100.0%;	Pred. No. 1.le-35;	Length 734;
Matches 199;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;
QY	1	GTTTATTCCTCCACAGATCCCGCAGACAGACCTAGTCTGATGAAGCTGCTCCTTTG	60
Db	201	GTTTATTCCTCCACAGATCCCGCAGACAGACCTAGTCTGATGAAGCTGCTCCTTTG	142
QY	61	TCCTAATATTCATATTCACAGACAGACCAATTCCTGGCATTCACATTTTAAATTAATGTGGAA	120

```

Db      1.1  TCCATATTCATATTCACAGACACCATTCCTGTGCATTCACATTTTAAATATATGGAA  82
QY      1.1  GGGGATAGAGAGAACTGCGACACTGTCAATACGCTTAGGGCTGAATTTTGTACAGATAAATAA  180
Db      1.1  GGGGATAGAGAGAACTGCGACACTGTCAATACGCTTAGGGCTGAATTTTGTACAGATAAATAA  22
QY      1.1  ATAAATCATTCATCCCTTT 199
Db      1.1  ATAAATCATTCATCCCTTT 3
RESULT  1.1
LOCUS   A1721275.1
DEFINITION as82h08.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
IMAGE:2335263.3 similar to SW:ECIC.BOVIN P34281 EPITHELIAL
CHLORIDE CHANNEL PROTEIN ; mRNA sequence.
ACCESSION A1721275
VERSION   A1721275.1 GI:5038531
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 501)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gaisel,G., Jost,S.,
Kilman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Martin,M., Martin
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST project
Unpublished (1997)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
TEL: 314 286 1800
FAX: 314 286 1810
EMAIL: est@watson.wustl.edu
This clone is available royalty-free through LBNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -400P from Glibco
High quality sequence stop: 394.
LOCATION/Qualifiers
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2335263"
/clone_id="Barstead colon HPLRB7"
/sex="male"
/dev_host="adult, age 25"
/lab_host="DH10B (phage resistant)"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTGACATCTGAGAGGAGCGCGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[5' ATTACTACTAAT 3' and 5' ATTACTAGTG 3'], digested
with Not I and cloned into the Not I and Eco RI sites of
the modified pT733 vector. Library constructed by Bob
Barstead."
BASE COUNT 158 a 83 c 106 g 154 t
ORIGIN
Query Match 80.4%; Score 194.6; DB 10; Length 501;
Best Local Similarity 95.7%; Pred. No. 1.2e-34;
Matches 200; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
1 GTTATTCCTCCACAGACTCCGCCGAGACACACTGCTGATGAAAGCTGCTCCTTG 60
214 GTTATTCCTCCACAGACTCCGCCGAGACACACTGCTGATGAAAGCTGCTCCTTG 155

```

Cy 1 GTTATTCCACAGACTCCCACAGACACCTAGTCTGATGAACGTCTGCTCTTG 60
 |||||
Ddb 214 GTTATTCCACAGACTCGCACAGACACCTAGTCTGATGAACGTTGCTCTTG 155
 |||||

OY	61	TCCATTTTTCATTATCAACAGACCATTCCTGCGATTCACATTTTAAAAATATGTGGAA	120
Db	154	TCCATTTTTCATTATCAACAGACCATTCCTGCGATTCACATTTTAAAAATATGTGGAA	95
OY	121	GTCGATAGAGAACATGCAGCTGTCTCAATAAGCCTTAGGCGTGTAATTTTTTGTCCAGATAAATAA	180
Db	94	GTGCGATAGAGAACATGCAGCTGTCTCAATAAGCCTTAGGCGCTCACATTTTGTGCCATTAATAC	35
OY	181	ATTAATCATTCATTCCTTTTATTTTGTACTTAT 209	
Db	34	AAAAATCATTCATTCCTTTTATTTTATTTAT 6	
RESULT 12			
A1721121/c			
LOCUS			
DEFINITION	A1721121	335 bp	mRNA EST 10-JUN-1999
ACCESSION	a833d08.x1	Bartstead colon HPLRB7 Homo sapiens CDNA clone	
VERSION	IMAGE:2334351.3	' similar to TR:088826 088826 GDB-5 PROTEIN.	; MRNAN
KEYWORDS	sequence.		
SOURCE	A1721121		
ORGANISM	A1721121.1	GI:5038377	
REFERENCE	EST.		
AUTHORS	human.		
TITLE	Homo sapiens		
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
COMMENT	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1 (bases 1 to 335)		
	Hallier,L., Allen,M., Bowles,L., Dubque,T., Geisel,G., Jost,S.,		
	Krizman,D., Kucaba,T., Lacey,M., Le,N., Lennon,G., Marra,M., Martin		
	,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,		
	White,Y., Wylie,T., Waterston,R. and Wilson,R.		
	WashU-NCI Human EST Project		
	Unpublished (1997)		
	Contact: Wilson RK		
	Washington University School of Medicine		
	444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: est@watson.wustl.edu		
	This clone is available royalty-free through LNL ; contact the		
	IMAGE Consortium (info@image.lnl.gov) for further information.		
	Possible reversed clone: similarity on wrong strand		
	Seq primer: -40UP from Gibco		
	High quality sequence stop: 312.		
FEATURES	Location/Qualifiers		
source	1..335		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2334351"		
	/clone_id="Bartstead colon HPLRB7"		
	/sex="male"		
	/dev_stage="adult, age 25"		
	/lab_host="DH10B (phage resistant)"		
	/note="Organ: colon; Vector: pRT73D-Pac (Pharmacia) with a		
	modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st		
	strand cDNA was primed with a Not I - oligo(dT) primer [5'		
	TGTACCAATCTGAAGTGGAGCGCGCCGCCCCCTTTTATTTTATTTTATTTT		
	3']; double-stranded cDNA was ligated to Eco RI adaptors		
	[5' AATCACTAGTAAT 3' and 5' ATTACAGTG 3'] , digested		
	with Not I and cloned into the Not I and Eco RI sites of		
	the modified pRT73 vector. Library constructed by Bob		
	Bartstead."		
BASE COUNT	98 a	57 c	71 g 109 t
ORIGIN			
Query Match	76.0%;	Score 183.8;	DB 10; Length 335;
Best Local Similarity	96.4%;	Pred. No. 3.6e-32;	
Matches 188:	Conservative 0;	Mismatches 7;	Indels 0;
	Gaps 0;		
OY	1	GTTATTTCTCCACAGACTCGCGCCAGAGACACCTAGTCTCTGATGGAAGCTGCTCCTTG 60	
Db	195	GTTATTTCTCCACAGACTCGCGCCAGAGACACCTAGTCTCTGATGGAAGCTGCTCCTTG 136	

QY	1	TCCTAATTTTCATATTCACAGACGACCATTCCTGGCAATTCACATTTTAAAAATATATGCGAA	120
DB	135	TCCTAATTTTCATATTCACAGACGACCATTCCTGGCAATTCACATTTTAAAAATATATGCGAA	76
QY	141	GTGCATTTAGACAGAAACGACAGCTGTCTCAATATAGCCATGAGGCTGCAATTTTGTGCACATTAATAA	180
DB	75	GTGCATTTAGACAGAAACGACAGCTGTCTCAATATAGCCATGAGGCAATTTTGTGCACATTAATAA	16
QY	141	ATAAATTCATTCATTC 195	
DB	15	ATAAATTCATTCATTC 1	
RESULT	13		
LOCUS	AA296955	378 bp	EST
DEFINITION	EST112726	Colón I Homo sapiens cDNA 5' end, mRNA sequence.	18-APR-1997
ACCESSION	AA296955		
VERSION	AA296955.1	GI:1949515	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 378)		
	Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulmer,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,C., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritcham,J.L., Geoghegan,N.S., Glodde,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,J.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudak,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., DiMke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,M.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.		
TITLE:	Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence		
JOURNAL MEDLINE	Nature 377 (6547 Suppl), 3-174 (1995)		
COMMENT:	96026280		
	Other_ESTs: THC167738		
	Contact: Kerlavage, AR		
	Bioinformatics		
	The Institute for Genomic Research		
	9712 Medical Center Drive, Rockville, MD 20850 USA		
	Tel: 3018699056		
	Fax: 3018699423		
	Email: arkerlav@tigr.org		
	For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/tdb/hgi/hgi.html)		
	Seq primer: M13 Reverse.		
FEATURES	Location/Qualifiers		
Source	1..378		
	/organism="Homo sapiens"		
	/db_xref="ATCC (lnhost):1.27834"		
	/db_xref="taxon:9606"		
	/clone_lib="Colón I"		
	/dev_stage="adult"		
	/note="Organ: Colón; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT	112 a 79 c 64 g 119 t 4 others		
CRIGIN			
Query Match	74.88; Score 181; DB 10; Length 378;		

Best Local Similarity 94.7%; Pred. No. 1.5e-31;
Matches 196; Conservative 0; Mismatches 9; Indels 2; Gaps 1

29 1 GTTATTCTCCACAGACTCCGACAGACACCTAGTCTGATGAACGTCGTCCITG 60

Db 174 GTTATTCCTCCACAGACTCCGCCAGAGACACCTAGTCCCTGATGAACGCTCTCCCTTG 233

61 TCCTAATATTCATATCAACAGCACCATTCCCTGGCATTCACATTTTAAAAATTATGGAA 120

Db 234 TCCATATTCATACACGACCAATTCGCAATCACATTTAAAAATATGGA 293

07 121 GTGGATAGGAGAACCTGCAGCTGTCAATACCCATGGCCGAAATTTGGCAGATAAAATATAT

Db 294 GTGGATAGGAGAACCTGCAGCTGTCAATAGNCIAGGGCTGAATTTTGGCAGGTTAAATATA 33

181 ATAAATCAITTCATCCCTTTTGGATT 207

DB 354 --1AA1CATTTCANCN111111GAT 3/8

RESULT 14

LOCUS	161 bp	MRNA	EST	06-APR
AI582072				

IMAGE: 2173436 3', mRNA sequence.

ACCESSION	AI582072
VERSION	AI582072.1
	GI:4567969

ALTERNATES	201:
SOURCE	human.

Eukaryota; Metazoa; Ch

REFERENCE 1 (bases 1 to 161)

Krizman, D., Kucaba, T.,

White, Y., Wylie, T., W...


```
OY 21 CGCCAGAGACACCTAGTCTGATGAAAGCTGCTGCTGCTGCTAAATATTCATATCAACA 80
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 102 CTCAGAGAGCCGCCATTCGCCGAAAGACTCAACTCCCTTGCTGACATCAGCATCAACA 161
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 81 GCACCATTCCTGGCATTCACATTTTAAAAATTTGTGGAAGTGATAGAGAACTGCAGC 140
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 162 GCACCATTCCTGGCATTCACAGTGTGTAAGTAATGTGGAAGTGCTAGGGGAAATGCAGG 221
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 141 TGTCAATAGCCCTAGGCGTGAATTTTGTGATATAATTAATATCATTCATCCTTTT 200
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 222 TGACACTAGGTTTGCACTGATTTTTCAGGCAAGAAATCAACCA-----GTGATTCT 273
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 201 TTTGATTATATAAATTTCTAAAAATGATTTTATAGACTTCCTGT 242
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 274 TTCACCTGAGAAATTTCTAAAAATGTACTTATAGACTTCCTGT 315
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: April 3, 2002, 20:27:47
Job time: 40230 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:54:41 ; Search time 521.98 Seconds
(without alignments)
105,000 Million cell updates/sec

Title: US-09-049-696-16

Perfect score: 242
Sequence: 1 GTTATTCCTCCACAGACATC.....ATGTTATTTAGACTTCCTGT 242

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_NA:*

1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCRTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/Backfillseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166.8	68.9	878	1	US-08-469-667-8
2	166.8	68.9	878	5	PCT-US95-07289-8
3	33.8	14.0	805	1	US-08-118-469A-6
4	33.8	14.0	805	1	US-08-909-119-6
5	33.2	13.7	2520	2	US-08-454-557C-50
6	33.2	13.7	2520	2	US-08-340-426D-50
7	33.2	13.7	2520	2	US-08-450-673C-50
8	33.2	13.7	2520	5	PCT-US95-1711A-50
9	33.2	13.7	10684	2	US-08-487-826B-13
10	32.6	13.5	19124	2	US-08-333-008A-7
11	31.6	13.1	6519	4	US-09-328-111-731
12	30.6	12.6	639	3	US-08-714-918-16
13	30.6	12.6	2018	4	US-09-265-315-16
14	30.6	12.6	2018	4	US-09-265-315-16
15	30.6	12.6	2018	4	US-09-265-315-16
16	30.6	12.6	2018	4	US-09-265-315-16
17	30.6	12.6	2187	2	US-08-679-635A-1
18	30.6	12.6	4248	3	US-08-678-614-1
19	30.4	12.6	1395	1	US-07-991-867B-25
20	30.4	12.6	1395	1	US-08-107-755A-25
21	30.4	12.6	1395	2	US-08-544-332-25
22	30.4	12.6	5433	3	US-08-929-329-1
23	30.4	12.6	6768	1	US-08-107-755A-1
24	30.4	12.6	7488	3	US-08-475-886-3
25	30.4	12.6	8457	1	US-07-991-867B-1
26	30.4	12.6	8457	2	US-08-544-332-1
27	30.2	12.5	2757	1	US-08-459-415-2

C	28	30.2	12.5	2757	5	PCT-US96-06352-90	Sequence 90, Appl
C	29	30.2	12.5	2757	5	PCT-US96-06352-90	Sequence 90, Appl
C	30	30.2	12.5	7486	4	US-08-475-886-5	Sequence 5, Appl
C	31	30.2	12.5	7486	4	US-08-397-232-3	Sequence 3, Appl
C	32	30.2	12.5	7486	4	US-08-475-886-3	Sequence 3, Appl
C	33	30	12.4	1185	2	US-08-179-557-14	Sequence 14, Appl
C	34	30	12.4	4279	4	US-08-993-825-1	Sequence 1, Appl
C	35	29.8	12.3	1415	1	US-08-413-118-127	Sequence 127, App
C	36	29.8	12.3	1415	1	US-08-473-446-127	Sequence 13, Appl
C	37	29.8	12.3	19124	2	US-08-487-826B-13	Sequence 1, Appl
C	38	29.6	12.2	1221	1	US-08-445-090-1	Sequence 13, Appl
C	39	29.6	12.2	1221	4	US-09-286-691-13	Sequence 13, Appl
C	40	29.6	12.2	1221	4	US-09-687-147-13	Sequence 1, Appl
C	41	29.4	12.1	2861	1	US-08-299-953-1	Sequence 1, Appl
C	42	29.4	12.1	2861	1	US-08-459-415-1	Sequence 1, Appl
C	43	29.4	12.1	2861	5	PCT-US95-11231-1	Sequence 1, Appl
C	44	29.4	12.1	3881	1	US-08-299-953-2	Sequence 2, Appl
C	45	29.4	12.1	3881	1	US-08-459-415-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-459-667-8
Sequence 8, Application US/08469667
Patent No. 5733748

GENERAL INFORMATION:

APPLICANT: YU, Guo-Liang
APPLICANT: Rosen, Craig
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carrella, Byrne, Bain, Gilfillan, Cecchi,
ADDRESSER: Stewart & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ

COUNTRY: USA

ZIP: 07068-1739

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,667

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Ferraro, Gregory D.

REGISTRATION NUMBER: 36,134

TELEPHONE: 201-994-1744

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 878 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2..685

Query Match 68.9%; Score 166.8; DB 1; Length 878;

Best Local Similarity 91.6%; Pred. No. 3.7e-38;

Matches 196; Conservative 1; Mismatches 15; Indels 2; Gaps 2;

OY 1 GTTATTCCTCCACAGACTCGCCAGAGACACCTAGTCTGATGAAGCTGTCTCTTG 60
|||||
DB 625 GTTATTCCTCCACAGACTCGCCAGAGACACCTAGTCTGATGAAGCTGTCTCTTG 684
OY 61 T-CCTAATATTCATATNCAACACACCATCTCGGATTCACATTTTAAATATATGTGA 119
|||||
DB 665 TGGCTAATATTCATATNCAACACACCATCTCGGATTCACATTTTAAATATATGTGA 744
OY 120 AGTGATAGAGAGACTGCGAGCTGTCAATAGCTAGGCTGAATTTTGTACATTAATA 179
|||||
DB 745 AGTGGTAGGAGAGACTGCGAGCTGTCAATAGCTAGGCTGAATTTTGTGCGGTGAT-A 803
OY 180 AATAATCATTCATCCTTTTGTGATTAATAA 213
|||||
DB 804 AATAATCATTCATCCTTTTGTGATTAATAA 837

RESULT 2

PCT-US95-07289-8
Sequence 8, Application PC/TUS9507289
GENERAL INFORMATION:
APPLICANT: Yu, Guo-Liang
APPLICANT: Rosen, Craig
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
ADDRESSEE: Stewart & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07289
FILING DATE: 06-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-265
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
PCT-US95-07289-8

Query Match 68.9%; Score 166.8; DB 5; Length 878;
Best Local Similarity 91.6%; Pred. No. 3.7e-38;
Matches 196; Conservative 1; Mismatches 15; Indels 2; Gaps 2;

OY 1 GTTATTCCTCCACAGACTCGCCAGAGACACCTAGTCTGATGAAGCTGTCTCTTG 60
|||||
DB 625 GTTATTCCTCCACAGACTCGCCAGAGACACCTAGTCTGATGAAGCTGTCTCTTG 684
OY 61 T-CCTAATATTCATATNCAACACACCATCTCGGATTCACATTTTAAATATATGTGA 119
|||||
DB 665 TGGCTAATATTCATATNCAACACACCATCTCGGATTCACATTTTAAATATATGTGA 744

OY 120 AGTGATAGAGAGACTGCGAGCTGTCAATAGCTAGGCTGAATTTTGTACATTAATA 179
|||||
DB 745 AGTGGTAGGAGAGACTGCGAGCTGTCAATAGCTAGGCTGAATTTTGTGCGGTGAT-A 803
OY 180 AATAATCATTCATCCTTTTGTGATTAATAA 213
|||||
DB 804 AATAATCATTCATCCTTTTGTGATTAATAA 837

RESULT 3

US-08-118-469A-6
Sequence 6, Application US/08118469A
Patent No. 5656451
GENERAL INFORMATION:
APPLICANT: Flavell, Richard A.
APPLICANT: Fikrig, Erol
APPLICANT: Lam, Tuan T.
APPLICANT: Kantor, Fred S.
APPLICANT: Barthold, Stephen W.
TITLE OF INVENTION: NOVEL B. BURKODORFERI POLYPEPTIDES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: C/O FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,469A
FILING DATE: 08-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,757
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: YU-102CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 805 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 130..711
US-08-118-469A-6

Query Match 14.0%; Score 33.8; DB 1; Length 805;
Best Local Similarity 49.7%; Pred. No. 0.72; Mismatches 87; Indels 0; Gaps 0;
Matches 86; Conservative 0;

OY 70 TCATATCAACAGACACCATCTCGGATTCACATTTTAAATATATGTGAAGTGAAG 129
|||||
DB 1 TCATATCAACAGACACCATCTCGGATTCACATTTTAAATATATGTGAAGTGAAG 60
OY 130 AGACTGACGCTGTCAATAGCTAGGCTGAATTTTGTCAATTAATAATATATCAT 189
|||||

Db 61 AATAATATTTATTTGTAACCTTACTTTTAAATTTAATGATTAATAATTAAGG 120

QY 190 TCATCCTTTTGGATTATTAATAATTTCTAATAATGATTTAGACTCCTGT 242

Db 121 AGAATTTTATGATTAATAATGATTTTAAACTATTTGTCATTCCTTT 173

RESULT 4

US-08-909-119-6
Sequence 6, Application US/08909119
Patent No. 5807685

GENERAL INFORMATION:
APPLICANT: Flavell, Richard A.
APPLICANT: Fikrig, Erol
APPLICANT: Lam, Tuan T.
APPLICANT: Kantor, Fred S.
APPLICANT: Barthold, Stephen W.
TITLE OF INVENTION: NOVEL B. BURGONFERRI POLYPEPTIDES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: C/O FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/909,119
FILING DATE: 11-AUG-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/118,469
FILING DATE: 08-SEP-1993
APPLICATION NUMBER: US 08/099,757
FILING DATE: 30-JUL-1993

ATTORNEY/AGENT INFORMATION:
NAME: Haley JT., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: YU-102CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 805 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO

FEATURE:
NAME/KEY: CDS
LOCATION: 130..711
US-08-909-119-6

Query Match 14.0%; Score 33.8; DB 1; Length 805;
Best Local Similarity 49.7%; Pred. No. 0.72;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 70 TCATTCACACACACACATTCCTGGCATTTTCAATTTTAAAAATTTATGTGAGTGAAGG 129

Db 1 TCATTTATTAAGACCTCCTGTTTCATTTTAACTATTTGTTTAAAGTGCTACA 60

QY 130 ACAACTGCAGCTGCTCAATGAGGCTGAGATTTTTCAGATTAATAAATCAAT 189

Db 61 AATAATATTTATTTATTTGTAACCTTACTTTTAAATTTAATGATTAATAATTAAGG 120

QY 190 TCATCCTTTTGGATTATTAATAATTTCTAATAATGATTTAGACTCCTGT 242

Db 121 AGAATTTTATGATTAATAATGATTTTAAACTATTTGTCATTCCTTT 173

RESULT 5

US-08-454-557C-50/c
Sequence 50, Application US/08454557C
Patent No. 5830670

GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Mands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-50

Query Match 13.7%; Score 33.2; DB 2; Length 2520;
Best Local Similarity 57.8%; Pred. No. 1.5;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 93 GCATTACATTTTAAAAATTTATGTGAGTGAAGAGTGCAGCTGTCAATAGCCT 152

Db 2412 GAACACATGCTTTAAATATGACAGTGAAGAGGAGGCTGAATTAACACCAACAAGCTA 2353

QY 153 AGGCTGATTTTTCGACATTAATAATTAATCAATCATC 194

Db 2352 AAAAAAGAAATCTTGATTGATTATTAATTAACAAAGCATGATC 2311

RESULT 6

US-08-340-426D-50/c
Sequence 50, Application US/08340426D
Patent No. 5948634

GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Mands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

REFERENCE/DOCKET NUMBER: 0609.3840004

PCF-0595-1/11

Query Match	13.7%	Score 33.2	DB 5	Length 2520
Best Local Similarity	57.8%	Pred. No. 1.5		
Matches 59	Conservative	0	Mismatches 43	Indels 0
				Gaps 0

Db 3834 AAAATA 3840

RESULT 11

US-08-233-008A-7/c
Sequence 7, Application US/08233008A
Patent No. 5578480

GENERAL INFORMATION:

APPLICANT: Khandke, Kiran M.
TITLE OF INVENTION: Methods For The Isolation And
Purification Of The Recombinantly Expressed chondroitinase
TITLE OF INVENTION: I and II Enzymes From P. vulgaris
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07470-8426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,008A

FILING DATE: 22-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31,885-01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-831-3244

TELEFAX: 201-831-3305

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 6519 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:NAME/KEY: CDS
LOCATION: 3238..6276

US-08-233-008A-7

Query Match 13.1%; Score 31.6; DB 1; Length 6519;
Best Local Similarity 58.5%; Pred No. 5.8; Mismatches 39; Indels 0; Gaps 0;

Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 141 TGTCAATAGCCAGGCGATTTTGTGAGATAATAAATCAATTCCTTTT 200

DB 6414 TATCATATAAAGCGATTCATTTTATTTTCAAAATTAATTCATTCCTTAA 6355

QY 201 TTGATATATAAATTTCTAATAATGATTTTGA 234

DB 6354 CTGTGTTTAATTTTAAATAAAGAGCTCGA 6321

RESULT 12

US-09-328-111-731/c
Sequence 731, Application US/09328111
Patent No. 6262333

GENERAL INFORMATION:

APPLICANT: Endege, Wilson O.
APPLICANT: Steimann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.

APPLICANT: Bushnell, Steven E.

APPLICANT: Carroll III, Eddie

APPLICANT: Carino, Theodore J.

APPLICANT: Dertl, Adnan

APPLICANT: Ford, Donna M.

APPLICANT: Lewis, Marcia E.

APPLICANT: Monahan, John E.

APPLICANT: Schlegel, Robert

TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

TITLE OF INVENTION: PRODUCTS

FILE/REFERENCE: CCD-257 (US)

CURRENT APPLICATION NUMBER: US/09/328,111

CURRENT FILING DATE: 1999-06-08

EARLIER APPLICATION NUMBER: US 60/088,801

EARLIER FILING DATE: 1998-06-10

NUMBER OF SEQ ID NOS: 850

SOFTWARE: PASTSEQ for Windows Version 3.0

SEQ ID NO 731

LENGTH: 639

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)...(639)

OTHER INFORMATION: n = A,T,C or G

US-09-328-111-731

Query Match 12.6%; Score 30.6; DB 4; Length 639;

Best Local Similarity 50.3%; Pred No. 5.3; Mismatches 74; Indels 0; Gaps 0;

Matches 75; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 83 ACCATTCCTGCGATTCACATTTTAAATAATATGGAAGTGATAGGAACTGCAGCTG 142

DB 334 ACAATTCCTTAATAGCCCTCTAGGTTTGAGGGAAGGAGATGAATGTGACATG 275

QY 113 TCAATAGCCCTGAGGCTGATTTTGTGAGATRAATAAATAATATATTCATCTTTT 202

DB 274 TGCACCTCACAAAGCGAGATTTCTTCTATATATAAAGTCACAGGACCTTTACTA 215

QY 203 TGATTATAAATTTCTAATAATGATTTT 231

DB 214 TGTGTTGGCATAGCCAGATTTTATCT 186

RESULT 13

US-08-714-918-16
Sequence 16, Application US/08714918
Patent No. 6037123

GENERAL INFORMATION:

APPLICANT: Benton, Bret
APPLICANT: Lee, Ving
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Flitch Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1

RESULT 1
US-09-265-315-16
Sequence 16, Application US/09265315
Patent No. 6187541
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmidt, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
NUMBER OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE:
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,315
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

RESULT: 115
 US-09-265-315-16
 Sequence 16, Application US/09265315
 Patent No. 6187541
 GENERAL INFORMATION:
 APPLICANT: Benton, Bret
 APPLICANT: Iee, Vang J.
 APPLICANT: Malouin, Francois
 APPLICANT: Martin, Patrick K.
 APPLICANT: Schmid, Molly B.
 APPLICANT: Sun, Dongxu
 TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
 TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
 TARGET GENES
 NUMBER OF INVENTION: 111
 NUMBER OF SEQUENCES: 111
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 MB
 MEDIUM TYPE: storage
 COMPUTER: IBM compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/265, 315
 FILING DATE: March 9, 1999
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/714, 918
 FILING DATE: September 13, 1996
 APPLICATION NUMBER: 60/009,102

FILING DATE: December 22, 1995
 APPLICATION NUMBER: 60/003,798
 FILING DATE: September 15, 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 240/247
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2018 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-265-315-16

Query Match 12.6%; Score 30.6; DB 4; Length 2018;
 Best Local Similarity 58.1%; Pred. No. 7.6;
 Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 130 AGAAGTGCAGCTGTCATATAGCGCTAGGCTGATTTTGTGACATAATAATAATCAT 189
 DB 943 AAAATACCGACATTAATTTCTGATACATAATTTTGTGCTGATTAATTCATTAAG 1002
 QY 190 TCATCCTTTTGTGATTAATAAATTTCTTAA 222
 DB 1003 TCATCCTCTGTTGATTAATTCGCTGTTAA 1035

Search completed: April 3, 2002, 20:54:48
 Job time: 41846 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:18:25 ; Search time 1321.64 Seconds
(without alignments)
156,981 Million cell updates/sec

Title: US-09-049-696-16

Perfect score: 242
Sequence: 1 GTTATTCCTCCACAGACTC.....ATGATTTTAGACTTCCTGT 242

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N.Geneseq.1101.*

1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	242	100.0	3109	22	AAH35019 Human colon cancer
2	242	100.0	3111	20	AAZ09840 Human membrane spa
3	230	95.0	3311	22	AAI29502 C9orf2 determined c
4	213.4	88.2	2854	22	AAH34679 Human colon cancer
5	213.4	88.2	2854	22	AAH34679 Human colon cancer
6	213.4	88.2	2854	22	AAH34679 Human colon cancer
7	202.4	83.6	2825	22	AAH46124 Human secreted pro
8	201.2	83.1	255	16	AAH46124 Human colon cancer
9	166.8	68.9	878	18	AAAT5884 Human gene signatu
10	166.8	68.9	878	18	AAAT5884 Human colon specif
11	154	63.6	2745	20	AAV16672 Polynucleotide seq
					Human ICACC-1 nucl

ALIGNMENTS

12	151	62.4	2742	22	AAH46102	Human CLCA1 coding
13	92	38.0	2843	22	AAH46120	Mouse Gob-5 cDNA,
14	92	38.0	2931	20	AAH81925	Murine ICACC-1 nuc
15	74.8	30.9	2739	22	AAH46101	Mouse Gob-5 coding
16	50	20.7	936	22	AAH58252	Oligonucleotide D1
17	50	20.7	936	22	AAH58254	Oligonucleotide D1
18	50	20.7	936	22	AAH58257	Oligonucleotide D1
19	50	20.7	936	22	AAH58259	Oligonucleotide D1
20	50	20.7	936	22	AAH58262	Oligonucleotide D2
21	50	20.7	936	22	AAH58265	Oligonucleotide D2
22	48.4	20.0	936	22	AAH58262	Oligonucleotide D1
23	48.4	20.0	936	22	AAH58264	Oligonucleotide D1
24	48.4	20.0	936	22	AAH58265	Oligonucleotide D1
25	48.4	20.0	936	22	AAH58269	Oligonucleotide D2
26	48.4	20.0	936	22	AAH58262	Oligonucleotide D2
27	48.4	20.0	936	22	AAH58265	Oligonucleotide D1
28	48.4	19.8	244	22	AAH58268	Oligonucleotide D1
29	46.8	19.3	244	22	AAH58268	Oligonucleotide D1
30	40.2	16.6	61150	21	AAH58238	Oligonucleotide D1
31	37.8	15.6	8622	22	AAH58432	Murine MOCIL genom
32	35.8	14.8	2723	20	AAH58432	Aradipops thalia
33	35.8	14.8	7235	20	AAH58432	Genomic DNA encodi
34	34.8	14.4	2483	22	AAH15185	Human cDNA sequenc
35	34.4	14.2	345	22	AAH15185	Probe #7977 used t
36	34.4	14.2	116277	20	AAH15185	Borrelia burgdorfe
37	34.4	14.2	134525	11	AAH20249	Total base sequenc
38	34.4	14.2	910715	20	AAH20248	Borrelia burgdorfe
39	34.2	14.1	2124	21	AAH80534	Human secreted pro
40	33.8	14.0	805	16	AAH83839	B. burgdorferi T5
41	33.6	13.9	1803	21	AAH26963	Human coenzyme A-u
42	33.6	13.9	2074	22	AAH26963	N. tabacum arginin
43	33.6	13.9	3501	21	AAH70202	Plasmodium falcipa
44	33.6	13.9	4321	22	AAH70202	Nicotiana tabacum
45	33.6	13.9	5454	21	AAH70236	Plasmodium falcipa

RESULTS

1	AAH35019	standard; cDNA; 3109 bp.
2	AAH35019	
3	AAH35019	
4	AAH35019	
5	AAH35019	
6	AAH35019	
7	AAH35019	
8	AAH35019	
9	AAH35019	
10	AAH35019	
11	AAH35019	
12	AAH35019	
13	AAH35019	
14	AAH35019	
15	AAH35019	
16	AAH35019	
17	AAH35019	
18	AAH35019	
19	AAH35019	
20	AAH35019	
21	AAH35019	
22	AAH35019	
23	AAH35019	
24	AAH35019	
25	AAH35019	
26	AAH35019	
27	AAH35019	
28	AAH35019	
29	AAH35019	
30	AAH35019	
31	AAH35019	
32	AAH35019	
33	AAH35019	
34	AAH35019	
35	AAH35019	
36	AAH35019	
37	AAH35019	
38	AAH35019	
39	AAH35019	
40	AAH35019	
41	AAH35019	
42	AAH35019	
43	AAH35019	
44	AAH35019	
45	AAH35019	

Human colon cancer antigen encoding cDNA seq ID NO: 2101.

Human colon cancer; colon cancer antigen; diagnosis; detection;

colorectal carcinoma; ss.

Human sapiens.

WC300122920-A2.

05-APR-2001.

28-SEP-2000; 2000WO-US26524.

29-SEP-1999; 99US-0157137.

03-NOV-1999; 99US-0163280.

(HUMA-) HUMAH GENOME SCI INC.

Ruben SM, Barash SC, Birse CE, Rosen CA;

WPT; 2001-235357/24.

P-PSDB; AAG75614.

Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers -

PS Claim 1: Page 3587-3588; 9803pp; English.
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAG77789 represent sequences used in the exemplification of the
 CC present invention. 682 and page 7053 of the sequence listing were
 CC N.B. Pages 666 to 7052 of the publication, meaning no sequences are present for
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

SO Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match 100.0%; Score 242; DB 22; Length 3109;
 Best Local Similarity 100.0%; Pred. No. 1.9e-52;
 Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTTATTCTCCACAGACTCCGCGAGACACCTAGTCTGTGATGAACGTGCTCCTTG 60
 Db 2479 gttattctccacagactccgcgagacacctagctgtgataaacgtctccttg 2538
 OY 61 TCCTAATATTCATATCAACAGCAGCATTCCTGCATTCACATTTTAAATATGAGAA 120
 Db 2539 tcctaataatcatatcaaacagcagcatctcctgcattcacaatttaaaatcagagaa 2598
 OY 121 GTGGATAGGAGAACTGCAATAGCCTAGGGCTGAAATTTTGTGATGAATATAA 180
 Db 2599 gtggatagagaaactgcaatagcctagggctgaattttgtgatgaatataa 2658
 OY 181 ATAAATCATTCATCTTTTGTGATTAATAATTTCTAAATGATTTTACACTTCT 240
 Db 2659 ataatacatctcatcttttgtgattataaattttctaataatgattttagacttcc 2718
 OY 241 GT 242
 Db 2719 gt 2720

RESULT 2
 ID AA209840 standard; cDNA; 3111 BP.
 AC AA209840;
 DT 26-NOV-1999 (first entry)
 DE Human membrane spanning protein MSP-5 cDNA fragment 2.
 KW Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;
 KW neoplastic disorder; immunological disorder; reproductive disorder;
 KW MSP-5; ds.
 OS Homo sapiens.
 PN W09946380-A2.
 PD 16-SEP-1999.
 PF 09-MAR-1999; 99WO-US05073.
 PR 13-MAR-1998; 98US-0039064.

XX (INCY-) INCYTE PHARM INC.
 PA Tang YT, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;
 PI Guegler KJ, Kaser MR, Baughn MR, Shah P;
 DR WFL: 1999-551409/46.
 DR P-PSDB; AAY33298.
 PT New human membrane spanning proteins used to, e.g. prevent and treat
 PT neoplastic disorders -
 PS Example 1: Page 80-81; 81pp; English.
 CC This invention describes novel human membrane spanning proteins (MSPs),
 CC and the polynucleotides encoding them. The products of the invention are
 CC used to diagnose, prevent and treat neoplastic, immunological and
 CC reproductive disorders. This sequence encodes a human membrane spanning
 CC protein MSP-5 fragment.

SO Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match 100.0%; Score 242; DB 20; Length 3111;
 Best Local Similarity 100.0%; Pred. No. 1.9e-52;
 Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTTATTCTCCACAGACTCCGCGAGACACCTAGTCTGTGATGAACGTGCTCCTTG 60
 Db 2635 gttattctccacagactccgcgagacacctagctgtgataaacgtctccttg 2684
 OY 61 TCCTAATATTCATATCAACAGCAGCATTCCTGCATTCACATTTTAAATATGAGAA 120
 Db 2685 tcctaataatcatatcaaacagcagcatctcctgcattcacaatttaaaatcagagaa 2744
 OY 121 GTGGATAGGAGAACTGCAATAGCCTAGGGCTGAAATTTTGTGATGAATATAA 180
 Db 2745 gtggatagagaaactgcaatagcctagggctgaattttgtgatgaatataa 2804
 OY 181 ATAAATCATTCATCTTTTGTGATTAATAATTTCTAAATGATTTTACACTTCT 240
 Db 2805 ataatacatctcatcttttgtgattataaattttctaataatgattttagacttcc 2864
 OY 241 GT 242
 Db 2865 gt 2866

RESULT 3
 ID AA129502 standard; cDNA; 3311 BP.
 AC AA129502;
 DT 12-OCT-2001 (first entry)
 DE C902P determined cDNA sequence.
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 KW immunogenic; gene therapy; vaccine; colonic cancer; ss.
 OS Homo sapiens.
 PN W0200149716-A2.
 PD 12-JUL-2001.
 PF 29-DEC-2000; 2000WO-US35596.
 PR 30-DEC-1999; 99US-0476296.
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.

19-MAY-2000; 2000US-0575251.
PR 29-JUN-2000; 2000US-0609448.
PR 28-AUG-2000; 2000US-0649811.
XX
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX WPI: 2001-441847/47.
DR
XX
XX Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer -
PS
XX
XX Claim 2; Page 425-426; 472pp; English.
XX
XX The present invention describes colon tumor associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of TCAPs by expressing inactive proteins or to supplement the
CC patients own production of them. Additionally, (II) may be used to
CC produce the TCAP proteins, by inserting the nucleic acids into a host
CC cell culturing the cell to express the protein. (II) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC polymerase chain reaction (PCR) and hybridisation assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. (I) may
CC also be used as antigens in the production of antibodies against TCAPs
CC and in assays to identify modulators of TCAP expression and activity.
CC Anti-(I) antibodies and antagonists may also be used to down regulate
CC TCAP expression and activity. The anti-(I) antibodies may also be used
CC as diagnostic agents for detecting the presence of TCAPs in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
CC and AA24494 to AA24523 represent nucleotide and amino acid sequences
CC given in the exemplification of the present invention.
XX
XX
XX Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;
SQ

Query Match 95.0%; Score 230; DB 22; Length 3311;
Best Local Similarity 99.6%; Pred. No. 2, 1e-49;
Matches 241; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GTTTATTCCTCCACAGACTCCGCCAGACACCTAGTCTGATGAAGCTGCTCCTTG 60
DB 2943 gttattctctccacagactccgccagagacactagctcgtatgaacgctcctctg 3002
QY 61 TCCTAATATTCATATCAAGACGACCTTCCTGCTGATCAATTTAAATTTGGGAA 120
DB 3003 tccatattcatatcaatacaagcaccatccctgcatcaccatttaaaattatcgtgaa 3062
QY 121 GTGATAGGAGAACTGACAGCTGTCAATAGCTAGGGCTGAATTTTGTGAGATAAATAA 180
DB 3063 gtagataggaagactcagctgcatcactagagcctgaattttgtcagaataataa 3122
QY 181 ATTAATCATTCATCCTTTTGTGATTAATAAATTTTCTAATAAATTTTGTAGCTTCT 240
DB 3123 ataataatcatcacc-ttttttgcattataaatttctaaatgtattttagaactcct 3181
QY 241 GT 242
DB 3182 gt 3183

RESULT 4
AAH34879
ID AAH34879 standard: cDNA; 2854 BP.
XX

AC AAH34879;
XX
XX
XX 03-SEP-2001 (first entry)
XX
XX
XX Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
XX
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 1; ss.
XX
XX Homo sapiens;
XX
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
XX 03-NOV-1999; 99US-0163280.
XX
XX (HUMAN-) HUMAN GENOME SCI INC.
XX
XX
XX Ruden SM, Barash SC, Birse CE, Rosen CA;
PI WPI: 2001-335357/24.
DR P: SDB; AAG75474.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
PS
XX
XX Claim 1; Page 3462-3463; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB7789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX
XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;
SQ

Query Match 88.2%; Score 213.4; DB 22; Length 2854;
Best Local Similarity 97.3%; Pred. No. 3, 4e-45;
Matches 217; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GTTTATTCCTCCACAGACTCCGCCAGACACCTAGTCTGATGAAGCTGCTCCTTG 60
DB 2626 gttattctctccacagactccgccagagacactagctcgtatgaacgctcctctg 2685
QY 61 TCCTAATATTCATATCAAGACGACCTTCCTGCTGATCAATTTAAATTTGGGAA 120
DB 2686 tccatattcatatcaatacaagcaccatccctgcatcaccatttaaaattatcgtgaa 2745
QY 121 GTGATAGGAGAACTGACAGCTGTCAATAGCTAGGGCTGAATTTTGTGAGATAAATAA 180
DB 2746 gtagataggaagactcagctgcatcactagagcctgaattttgtcagaataataa 2805
QY 181 ATTAATCATTCATCCTTTTGTGATTAATAAATTTTCTAATAA 223
DB 2805
XX

human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were obtained from 3'-directed cDNA libraries prepared from various human tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-untranslated sequence is unique to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library is constructed so as to reflect accurately the relative abundance of different mRNAs in the particular tissue from which it was derived. The appearance frequency of a given GS in a cDNA library can be determined (esp. using primers and probes derived from the GS sequences) as a means of diagnosing abnormal cell function or for recognising different cell types.

Sequence 255 BP; 76 A; 52 C; 39 G; 80 T; 8 other;

Query Match 83.1%; Score 201.2; DB 16; Length 255;
Best Local Similarity 95.8%; Pred. No. 2.5e-42;
Matches 203; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

2 TTTATTCCTCCACAGACTCCGCGAGACACCTAGTCTGATGAACGTCCTCTTGT 61
43 ttatncccccacagactccgcagagacactagtcctgataacgctcctctgt 102
62 CCTAATATTCATATCAACGACACCATTCCTGGCATTCACATTTTAAATATGGAAG 121
103 cctaatatcatatcacacagacacatcctcgtgcatcatttaaaatcaggaag 162
122 TGGATGAGAACTGCGAGCTGTCATAGCCTAGGGCTGAATTTTGTGAGATAAATAA 181
163 tggatgggaactgcagctgtaacatagcctagggctgaatttgcagntaantana 222
182 TAAATCATTCATCCCTTTTGTGATTAATAA 213
223 taatcatcatccttcttlttngnttataa 254

RESULT 9

AAT45884
ID AAT45884 standard; cDNA: 878 BP.

AC AAT45884;

DT 13-MAR-1997 (first entry)

DE Human colon specific gene CSG5 cDNA partial clone.

KW Colon specific gene; CSG5; colon cancer; metastasis; diagnosis;

KW gene therapy; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

FT CDS 2..691

FT /tag= a

PN MO9639419-A1.

PD 12-DEC-1996.

XX 06-JUN-1995; 95WO-US07289.

XX 06-JUN-1995; 95WO-US07289.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Yu G;

XX WPI; 1997-043054/04.

DR P-PSDB; AAM06548.

XX Human colon specific genes and their expression products - detection of which, in non-colon tissue samples, can be used as indication of

PI Colon cancer metastasis

XX Claim 1; Fig 5; 60pp; English.

XX 13-cDNA clones (AAT45880-92), most of them partial clones, correspond to human colon specific genes, designated CSG1, CSG2, etc., that are primarily expressed in tissues derived from the colon. CSG7 and CSG10 show reduced expression in colon cancer cells as compared to that in normal cells; the remaining genes are overexpressed in colon cancer. The partial cDNA sequences can be used to isolate full-length clones and genomic clones including the complete gene. CSG nucleic acids can be used to produce CSG polypeptides (see also AAM06545-53) in transformed host cells, as probes to detect disorders of the colon, partic. colon cancer and colon cancer metastasis, and in gene therapy.

XX Sequence 878 BP; 257 A; 179 C; 188 G; 241 T; 13 other;

Query Match 68.9%; Score 166.8; DB 18; Length 878;
Best Local Similarity 91.6%; Pred. No. 1.8e-33;
Matches 196; Conservative 1; Mismatches 15; Indels 2; Gaps 2;

1 GTTATTCCTCCACAGACTCCGCGAGACACCTAGTCTGATGAACGTCCTCTTGT 60
635 gtattcctccacagactccgcagagacactagtcctgataacgctcctctgt 684
61 T-CCTAATATTCATATCAACGACACCATTCCTGGCATTCACATTTTAAATATGGA 119
605 tgcctaatatcatatcacacagacacatcctcgtgcatcatttaaaatcagga 744
120 AGTGGATGAGAACTGCGAGCTGTCATAGCCTAGGGCTGAATTTTGTGAGATAAATA 179
745 agtggatgggaactgcagctgtaacatagcctagggctgaatttgcagntaantana 803
180 AATATCATTCATCCCTTTTGTGATTAATAA 213
804 aataatcatccttcttlttngnttataa 837

RESULT 10

AAV16672
ID AAV16672 standard; cDNA: 878 BP.

AC AAV16672;

DT 22-JUN-1998 (first entry)

DE Polynucleotide sequence of a colon-specific gene.

KW Colon-specific gene; probe; detection; expression; human;

KW diagnostic assay; colon cancer; antibody; screening; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

FT CDS 2..685

FT /tag= a

PN US5733748-A.

PD 31-MAR-1998.

XX 06-JUN-1995; 95US-0469667.

XX 06-JUN-1995; 95US-0469667.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen C, Yu G;

XX WPI; 1998-229823/20.

PN	XX	WO200138530-A1.
XX	XX	
PD	XX	31-MAY-2001.
PF	XX	22-NOV-2000; 2000WO-JP08232.
PR	XX	24-NOV-1999; 99JP-0333479.
PR	XX	27-APR-2000; 2000JP-0127589.
PA	XX	(TAKE) TAKEDA CHEM IND LTD.
P1	XX	Nakanishi A, Morita S.
DR	XX	WPI; 2001-355935/37.
DR	XX	P-PSDB; AAB73716.
PT	XX	New antisense nucleotide, useful for treatment and prevention of
PT	XX	bronchial asthma and chronic obstructive pulmonary disease -
PS	XX	Claim 3; Page 82-84; 104pp; Japanese.
CC	XX	The invention relates to an antisense nucleotide targetted to the mouse
CC	XX	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart.
CC	XX	The C1CA1 gene (coding sequence shown in AAH6102). The invention also
CC	XX	relates to an antibody specific for the Gob-5 protein, medical and
CC	XX	diagnostic compositions containing the antisense nucleotide or the
CC	XX	antibody, and methods and kits for screening for compounds which inhibit
CC	XX	the protein. Gob-5 and C1CA1 are proteins expressed by goblet cells.
CC	XX	The antisense oligonucleotides and antibody are therefore useful for the
CC	XX	treatment and prevention of bronchial asthma and chronic obstructive
CC	XX	pulmonary disease. The present sequence represents the human C1CA1
CC	XX	gene coding sequence.
CC	XX	
CC	XX	Sequence 2742 BP; 833 A; 616 C; 623 G; 670 T; 0 other;

		62.4%;	Score 151;	DB 22;	Length 2742;
		Query Match Similarity,	100.0%;	Pred. NO.	2.5e-29;
		Best Local Similarity,			
Matches	151;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0
Oy	1	GTATTATTCCTCCACAGACTCGCCGACAGACACCCTA	TCTTGATGAAGCTCTGCTCCTTG	60	
Db	2592	gttattcctccacagactcgccgacagacacctagt	cctgatgaacgctctgcctcttg	2651	
Oy	61	TGCTAATATTTCAATFACAAGACGACCATTTTCTGG	CATTTCACATTTTAAAAATTATGTGGAA	120	
Db	2652	tccataatcatcatcaacaagaccattctgycattca	atttaaaattatgtggaa	2711	
Oy	121	GTGGATAGGAGAAGTCGACAGCTGCATAATAGCC	151		
Db	2712	gtgatataggagaactgcagctctgccaatagcc	2742		
RESULT	13				
ID	AAH46120				
	AAH46120 standard; CDNA; 2843 BP.				
xx					
AC	AAH46120;				
xx					
DT	11-SEP-2001 (first entry)				
DE					
Mouse	Gob-5 cDNA, SEQ ID NO:22.				
xx					
KW	Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;				
RV	expression inhibition; antisense therapy; gene therapy; bronchial asthma;				
xx	chronic obstructive pulmonary disease; antiasthmatic; ss.				
OS					
Mus	sp.				
xx					
FH					
key	Location/Qualifiers				
CDS	15..2756				
FT	/tag= a				

FT XX /product= "Mouse Gob-5"
 EN PN MO200138530-AL.
 FD 31 MAY-2001.
 XX 22: NOV-2000; 2000MO-JF082232.
 XX 24: NOV-1999; 99JP-0333479.
 XX 27: APR-2000; 2000JP-0127589.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX
 XX Naikanishi A., Morita S;
 XX WPI: 2001-355935/37.
 XX P-PSDB; AAB73715.
 XX
 XX New antitense nucleotide, useful for treatment and prevention of
 FT bronchial asthma and chronic obstructive pulmonary disease
 PT
 XX
 XX Example 1; Page 89-91; 104pp; Japanese.
 XX
 XX The invention relates to an antitense nucleotide, targetted to the mouse
 CC Gob-5 gene (coding sequence shown in AAB46101) or its human counterpart,
 CC the CLCA1 gene (coding sequence shown in AAB46102). The invention also
 CC relates to an antibody specific for the Gob-5 protein, medical and
 CC diagnostic compositions containing the antitense nucleotide or the
 CC antibody, and methods and kits for screening for compounds which inhibit
 CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 CC Ths antitense oligonucleotides and antibody are therefore useful for the
 CC treatment and prevention of bronchial asthma and chronic obstructive
 CC pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
 XX
 XX Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;

	Query Match	Score 92;	DB 22;	Length 2843;
	Nucleic Acid Similarity	67.1%;	Pred No.	2.4e-14;
Batches	149;	Conservative	0;	Mismatches 65; Indels 8; Gaps 1
DY	21 GCACGAGACAACCTATGTCTTGATGAACGCCTGGTCCCTCATATTTCATATCAACA	80		
db	2617 ctccaggaacgcgcccaattcccagaaggctccaatcccccttgctcgtgcatcacataacca	2676		
CY	21 GCACCATTCCTCGCATTTCACATTTTAATAAATATATGTGCCAAAGTAGAGAGAACAAC	140		
Db	2677 gcaccatttccttggcactlccacagtgcgcgaagaataatlgaaagtcgycgaaggaatatcgac	2736		
OY	141 TGTCAAATGACCAGAGCGTGAATTTTTGTGCAGATTAATAAATAAACATTCATTCCTTTT	200		
Dy	2737 tgacctcgtgltttgcacacgaatttccagcaagaataaca-----gtcaattcct	2788		
OY	201 TTTCATTATPAAAATTTTCTAAAAATGATTTTAGACTTCCTGT	242		
Dy	2719 tttaactggagaatttctaanaaatgtacttagacttlcctgl	2830		
<div style="text-align:center;">RESULT_14 LAFB1925</div>				
ID	AAF81925 standard; cDNA:	2931 BP.		
XX	AAF81925;			
XX				
Xx	13-JUN-2001 (first entry)			
Xx				
xX	Murine ICAC-1 nucleotide sequence.			
xx				
IxCAC-1; ICAC-2; asthma; atopic allergy; asthma-related disorder; interleukin 9 induced calcium activated chloride channel; IL-9; calcium activated chloride channel; anti-allergic; anti-asthmatic; anti-inflammatory; immunomodulatory; cystic fibrosis; Inflammatory bowel disease; autoimmune disease; ss.				

XX	Mus sp.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	CDS	8..2749
FM		/tag= A
FT		/product= "ICACC-1"
XX		/note= "IL-9 induced calcium activated chloride channel"
PN	MO9944620-AI.	
XX		
FD	10-SEP-1999.	
XX		
PF	03-MAR-1999;	99WO-US04703.
XX		
PR	03-MAR-1998;	98US-0076815.
XX		
PA	(MAGA-) MAGAININ PHARM INC.	
XX		
P1	Holroyd KJ, Levitt RC, Maloy WL, Louahed J, McLane M;	
P1	Nicolaides NC, Zhou Y, Dong Q;	
XX		
DR	WPI: 1999-550979/46.	
DR	P-PDB; AAB74822.	
PT	New nucleic acid encoding calcium activated chloride channel, used to	
PT	identify, e.g. specific modulators for treating atopic allergy -	
XX		
PS	Claim 2; Fig 2; 75pp; English.	
XX		
CC	The present sequence encodes the murine interleukin 9 (IL-9) induced	
CC	calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins	
CC	have anti-allergic, anti-asthmatic, anti-inflammatory and	
CC	immunomodulatory activities. Compounds (A) that downregulate ICACC are	
CC	used to alleviate asthma (or more generally atopic allergy), while those	
CC	(B) that upregulate ICACC are used to treat cystic fibrosis (CF) and	
CC	inflammatory bowel disease (IBD) (or other autoimmune diseases).	
CC	Measuring levels of ICACC-1 can be used in the diagnosis of asthma	
CC	(increased levels) or IBD (reduced levels), also for monitoring	
CC	treatment of these conditions. The ICACC proteins can be used as:	
CC	(1) to raise specific antibodies (Ab), useful: (a) as immunoassay	
CC	reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as	
CC	(A) to disrupt bonding between ICACC and its ligands; and (iii) to	
CC	identify modulators and binding partners. ICACC polynucleotides can be	
CC	used to generate transgenic animals or recombinant cells, used to screen	
CC	for antagonists, also as a source of therapeutic antisense agents or	
CC	diagnostic probes (for quantifying mRNA expression, e.g. for	
CC	identification of modulators).	
XX		
S0	Sequence 2931 BP; 867 A; 719 C; 684 G; 661 T; 0 other:	
	Query Match	38.0%; Score 92; DB 20; Length 2931;
	Best Local Similarity	67.1%; Pred. No. 2.5e-14;
	Matches 149; Conservative 0; Mismatches 65; Indels 8; Gaps 1,	
OY	21 CGCGGACACACCTAGTCTCATGAAGCTGTGCCTTGTCCTTAATATTCATCAACA 80	
Db	2610 ctcaaggacgcgccattcccgaaagcactacccttcgttcgacacgaatcaaca 2669	
OY	81 GCACCATTCCTCGCATTWCACATTTTAAAAAATTTATGTGGAAAGTAGAGAACCTGCACC 140	
Db	2670 gacaccttcctggcatccaacgctgcgaagataatgtgaaagtgcctaggagaatgcagc 2729	
OY	141 TGTCATATGCTTAGGCGCTGAATTTTGTGCACGATAATAATAAATAATTCATTCCTTTT 200	
Db	2730 tgaacctagtgattgcactgaatttttccacgcaagaatacaacca-----gtcattcct 2781	
OY	201 TTTCATTAATAAATTTTCTAAAATGATATTTATTTAGACTTCCTGT 242	
Db	2782 ttcaactggaatatcttcaaaaatgtlacttltagacttctcgt 2823	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```


Thu Apr 4 09:27:33 2002

Db 2723 tgacactagtttg 2736

Search completed: April 4, 2002, 03:18:31
Job time: 60263 sec

us-09-049-696-15.rng

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:05:40 ; Search time 22700.8 Seconds
(Without alignments)
175.867 Million cell updates/sec

Title: US-09-049-696-16
Sequence: 1 GTTATTCCTCCACAGACTC.....ATGATTTTACACTTCTGT 242

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: 1: gb_ba: 2: gb_htg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_vl: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_om: 20: em_ov: 21: em_pat: 22: em_ph: 23: em_pl: 24: em_ro: 25: em_sts: 26: em_sy: 27: em_un: 28: em_vl: 29: em_vl: 30: em_htgo_hum: 31: em_htgo_inv: 32: em_htgo_rod: 33: em_htg_hum: 34: em_htg_inv: 35: em_htg_rod: 36: em_htg_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	242	100.0	113764	9	HSJ651E10
2	242	100.0	164891	2	AL358950
3	230	95.0	3311	6	AX193489
4	230	95.0	3311	9	AF039400
5	228.4	94.4	35278	9	AF039401
6	213.4	88.2	2826	9	AF127036
7	197	81.4	2022	9	AK024970
8	166.8	68.9	878	6	I95746
9	125.8	52.0	3079	4	AF095584
10	92	38.0	2937	10	AB017156
11	43.8	18.1	256172	2	AC005139
12	43.8	18.1	310779	2	AC005140
13	43.4	17.9	178920	8	AL339195
14	41.8	17.3	103787	8	AT78B10
15	41	16.9	678	3	AF214670
16	41	16.9	2950	3	AFMPCOX
17	41	16.9	152409	2	AC018070
18	41	16.9	16552	2	AC018070
19	40.8	16.9	107702	8	ATF17L22
20	40.8	16.9	186989	3	AC007812
21	40.8	16.9	189117	3	AC009395
22	40.8	16.9	194916	8	ATCHRIV5
23	40.8	16.9	255624	3	AE003723
24	40.8	16.8	49885	6	AX059537
25	40.6	16.8	53778	8	TIJ1
26	40.6	16.8	176305	2	AC007943
27	40.6	16.8	198563	8	ATCHRIV14
28	40.6	16.6	45120	6	AX059511
29	40.2	16.6	84405	8	AC005693
30	40.2	16.5	1805	9	AB048893
31	40	16.5	201324	2	AC079637
32	40	16.5	117752	9	AC004743
33	39.6	16.3	1812	1	AF280099
34	39.4	16.3	110000	2	PFMALP1_0
35	39.4	16.3	155693	3	CEY39E4B
36	39.4	16.3	217686	2	AC015915
37	39.4	16.3	287612	2	AC091339
38	39.4	16.2	73345	8	AL0294725
39	39.2	16.2	169364	3	CEY56A3A
40	39.2	16.2	174019	9	AP001538
41	39.2	16.2	340000	9	AP001674
42	39.2	16.1	232	3	TRBVSGET2
43	39	16.1	275	3	TRBVSGET2
44	39	16.1	108365	8	AC015447
45	39	16.1	108365	8	AC015447

ALIGNMENTS

RESULT 1
LOCUS HSJ651E10
DEFINITION Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1, complete sequence.
ACCESSION AL122002
VERSION AL122002.16 GI:8247274
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 113764)
AUTHORS Brown/A.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT On Jun 4, 2000 this sequence version replaced gi:8247030. During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Ch1>

RP4-651E10 is from the library RPci-4 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pCYPAC2
This sequence is the entire insert of clone RP4-651E10.

FEATURES

source

1. 113764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p22.3-31.1"
/clone="RP4-651E10"
/clone_lib="RPci-4"

BASE COUNT 34890 a 21989 c 22351 g 34534 t
ORIGIN

Query Match

Best Local Similarity 100.0%; Score 242; DB 9; Length 113764;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTATTCCTCCAGAGACTCGCGAGAGACCTAGTCTGATGAAAGCTGCTCCTTG 60
D 32862 GTTATTCCTCCAGAGACTCGCGAGAGACCTAGTCTGATGAAAGCTGCTCCTTG 32921
QY 61 TCCAAATTCATATTCACAGACACCATTCCTGCGATTCACATTTAAATTAATGCGAA 120
D 32922 TCCAAATTCATATTCACAGACACCATTCCTGCGATTCACATTTAAATTAATGCGAA 32981
QY 121 GTGATAGGAGACCTGAGCTGTCAATAGCTGATGCTGATTTTGTGAGATAATATA 180
D 32982 GTGATAGGAGACCTGAGCTGTCAATAGCTGATGCTGATTTTGTGAGATAATATA 33041
QY 181 ATAAATCATTCATCTTTTGTGATTAATAATTTCTAAATGTAATTTAGACTCT 240
D 33042 ATAAATCATTCATCTTTTGTGATTAATAATTTCTAAATGTAATTTAGACTCT 33101
QY 241 CT 242
D 33102 GT 33103

RESULT 2

AL358950/c

LOCUS Homo sapiens chromosome 1 clone RP11-444C12, ** SEQUENCING IN
DEFINITION

ACCESSION

AL358950

VERSION

AL358950.4 GI:12539689

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT:

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 164891)
Plumb, B.
Direct Submission
Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jan 26, 2001 this sequence version replaced gi:9988471.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk
Project Information
Center project name: BA444C12
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least Q40
Consensus quality: 162610 bases at least Q30
Consensus quality: 163328 bases at least Q20
Insert size: 163691; sum-of-contigs
Insert size: 163373; 6.6% error; agarose-IP
Quality coverage: 5.19x in Q20 bases; sum-of-contigs quality
coverage: 5.21x in Q20 bases; agarose-IP

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 34236: contig of 34236 bp in length
* 34237 34336: gap of 100 bp
* 34337 47372: contig of 13036 bp in length
* 47373 47472: gap of 100 bp
* 47473 67025: contig of 19553 bp in length
* 67026 67125: gap of 100 bp
* 67126 71239: contig of 4114 bp in length
* 71240 71339: gap of 100 bp
* 71340 83903: contig of 12564 bp in length
* 83904 84003: gap of 100 bp
* 84004 91357: contig of 7534 bp in length
* 91358 91457: gap of 100 bp
* 91458 96452: contig of 4995 bp in length
* 96453 96552: gap of 100 bp
* 96553 148189: contig of 51637 bp in length
* 148190 148289: gap of 100 bp
* 148290 159778: contig of 11489 bp in length
* 159779 159878: gap of 100 bp
* 159879 164891: contig of 5013 bp in length.

FEATURES

source

1. 164891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-444C12"
/clone_lib="RPci-11.2"

misc_feature

1. 34236
/note="assembly fragment:00027
fragment_chain:1
clone_end:17
vector_side:left"

misc_feature

34337..47372
/note="assembly fragment:00312
fragment_chain:1"

misc_feature

47473..67025
/note="assembly fragment:01923
fragment_chain:1"

misc_feature

67126..71239

BASE COUNT	1028 a	692 c	742 g	849 t
ORIGIN				
Query=Match		95.0%;	Score 230;	DB 6;
Best Local Similarity		99.6%;	Pred. No. 2,4e-41;	
Matches 241;	Conservative	0;	Mismatches	0;
			Indels	1;
			Gaps	1;
QY	51	GTTTATTTCTCCACAGACTCCGCCACAGACACCTTAGTCTGTGATGAAGCTGCTCTTG	60	
DS	293	GTTTATTTCTCCACAGACTCCGCCACAGACACCTGTGCTGATGAAGCTGCTCTTG	3002	
OY	51	TCCATAATTCATATCAACAGACACCATTCCTCGGCACTTACATTTTAAAAATTAATGTGAA	120	
DO	3003	TCCATAATTTCTATATCAACAGACACCATTCCTCGGCACTTACATTTTAAAAATTAATGTGAA	3062	
OY	141	GTGCATAGAGAACTGCAGCTGTCTCAATAGACCTAGCGCTGAATTTTGTCCAGATAAATAA	180	
DB	3063	GTGCATAGAGAACTGCAGCTGTCTCAATAGACCTAGCGCTGAATTTTGTCCAGATAAATAA	3122	
OY	181	ATAAATTCATTCATCTTTTATTTTGTATTAATAAATTTTCTAAATGTATTTTAGACTTCT	240	
DO	3123	ATAAATTCATTCATCTCTTTTATTTTGTATTAATAAATTTTCTAAATGTATTTTAGACTTCT	3181	

DO	3182	GT 3183
RESULT	4	
LOCUS	AF039400	
DEFINITION	Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA,	
ACCESSION	AF039400	
VERSION	AF039400	
KEYWORDS	AF039400.1	GI:4009457
SOURCE	human:	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
	1 (bases 1 to 311)	
	Gruber,A.D., Eble,R.C., Ji,H.L., Schreuer,K.D., Fuller,C.M. and	
	Pallil,B.U.	
TITLE	Genomic cloning, molecular characterization, and functional	
	analysis of human CLCA1, the first human member of the family of	
JOURNAL	Ca2+-activated Cl- channel proteins	
MEDLINE	Genomics 54 (2), 200-214 (1998)	
	99047526	

```

TITLE: Direct Submission
JOURNAL: Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
FEATURES: Cornell University, Ithaca, NY 14853, USA
SOURCE: Location/Qualifiers
1..3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/tissue_type="small intestine"
1..3311
/gene="hclcal"
352..3096
/gene="hclcal"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPKSSVFLILHLLLEGALSNLQLNNNGYEGIVAIIDPNP
EDELLIOIKDWDYQASLIYFEATGKRFRFNVAAILPEFTWKADYVRPKLEYKNA
DIVIAESTPGNDPEPYTEQNGCNGEKERILHLPDFAGKKLAEYGGOGAGAFEMNH

```



```

LRMGVPEDEYNNDKEFYLSNGHIAVRCAGICTGTCTGATGGAAGCTGCTGCTCTG
VEKCEFEVLOSROTEKASIMFAQVHDSIVECTEQNNKRPKNOKCINRSTWEIT
RDESEFKRTPTMTPTPPTSLQIQRIYCLVDSGSMATGNRLNRINOAGQLFL
LQVLELSWGMVTFDSANAHQSELIOINGSDBDTLAKRLPAAASGSGTSCGLRBA
FTVIRKRYPTDGESEIVLLDGEDNTISGCFENYVSGAIIHTVALGSAOEMTELSK
MTGLOIYASDNOVONNGLIIDAFGALSSGNVSGRSIOLESKGLTIGNSOMMNTVIV
DSYVSKDTLFLITWTPTPOLILMDPSGOKGCGVVKRNMKMATVIOIPIJAKVGTMY
SVLALIESVNGKTYVLELIDNGAGADATKDGVSRYFTTYDNGRSVYKRALGVA
AARRVAPIPOOGSALYIPGMIEDEIOMNPPPELTKDDVOHKOVCFSRTSGSSEFVS
DVPNAPIPDLFPPGQITDILNAEIHGSLINLTWAPDDDHGTAKHYIIRISTIID
LRKFNESLQVNTTALIPKEANSEVEFLFKENITTFENGIDLFAIOAVKVDKSEI
SNIRVSLFIPQTPPTPPSPDETSAPOPNHINSTIPGIHILKIMKMWIGELQLSIA

BASE COUNT      1028 a      692 c      742 g      849 t
ORIGIN

Query Match      95.0%; Score 230; DB 9; Length 3311;
Best Local Similarity 93.6%; Pred. No. 2.4e-41;
Matches 241; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GTTATTCCTCCAGACAGCTCCGACAGACACCTAGTCTGATGGAAGCTGCTCTG 60
DB 2943 GTTATTCCTCCAGACAGCTCCGACAGACACCTAGTCTGATGGAAGCTGCTCTG 3002
OY 61 TCTTAATATTCATATCAGACAGACACCTGCGATTCATTTTAAATATATGCGAA 120
DB 3003 TCTTAATATTCATATCAGACAGACACCTGCGATTCATTTTAAATATATGCGAA 3062
OY 121 GGGATAGAGAGACAGCTGCTGATAGCTAGGCGTGAATTTTGTGCAGATTAATAA 180
DB 3063 GTGATAGAGAGACAGCTGCTGATAGCTAGGCGTGAATTTTGTGCAGATTAATAA 3122
OY 181 AATAATCATTCATCTTTTGTGATTAATAATTTTCTAAATGATTTTACACTTCT 240
DB 3123 AATAATCATTCATCTTTTGTGATTAATAATTTTCTAAATGATTTTACACTTCT 3181
OY 241 GT 242
DB 3182 GT 3183

RESULT 5
AF039401 35278 bp DNA PRI 14-DEC-1998
LOCUS AF039401 Homo sapiens calcium dependent chloride channel-1 (hclcal) gene,
DEFINITION complete cds.
ACCESSION AF039401.1 GI:4009459
VERSION AF039401
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 35278)
AUTHORS Gruber A.D., Elble R.C., Ji H.L., Schreier K.D., Fuller C.M. and
Pauli B.U.
TITLE Genomic cloning, molecular characterization, and functional
analysis of human Clcal, the first human member of the family of
Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 35278)
AUTHORS Gruber A.D., Elble R. and Pauli B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Department of Pathology, College of
Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES
source
1..35278
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"

```

```

/promoter
1..1617
/gene="hclcal"
1..33522
/gene="hclcal"
repeat_region
767..1265
/rpt_family="L1 transposable element"
/rpt_type="dispersed"
1587..1596
/gene="hclcal"
/exon
1618..1708
/note="putative"
/gene="hclcal"
/number=1
join(1618..1708,1962..2383,6667..6807,6979..7126,9696..9801,15454..15631,18591..18809,19783..20010,22293..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33310)
/gene="hclcal"
/product="calcium-dependent chloride channel-1"
1962..2383
/gene="hclcal"
/number=2
join(2222..2383,6667..6807,6979..7126,9696..9801,15454..15631,18591..18809,19783..20010,22253..24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33310)
/gene="hclcal"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/db_xref="GI:4009460"
/translation="MGPEKSSVFLILHLLEGALSNLSIQLNNNGEGIVAAIDPNVP
EDLEILIOIKDMVTOASLYLEATGKRFYKRNVALIPETWTKRADYVRKLEKNA
DVIAESTPPGNDPEPTOMGNGEGERHILPDPFJAGKKLAEYGOGAFFYEMAH
LRMGVPEDEYNNDKEFYLSNGHIAVRCAGICTGTCTGATGGAAGCTGCTCTG
VEKCEFEVLOSROTEKASIMFAQVHDSIVECTEQNNKRPKNOKCINRSTWEIT
RDESEFKRTPTMTPTPPTSLQIQRIYCLVDSGSMATGNRLNRINOAGQLFL
LQVLELSWGMVTFDSANAHQSELIOINGSDBDTLAKRLPAAASGSGTSCGLRBA
FTVIRKRYPTDGESEIVLLDGEDNTISGCFENYVSGAIIHTVALGSAOEMTELSK
MTGLOIYASDNOVONNGLIIDAFGALSSGNVSGRSIOLESKGLTIGNSOMMNTVIV
DSYVSKDTLFLITWTPTPOLILMDPSGOKGCGVVKRNMKMATVIOIPIJAKVGTMY
SVLALIESVNGKTYVLELIDNGAGADATKDGVSRYFTTYDNGRSVYKRALGVA
AARRVAPIPOOGSALYIPGMIEDEIOMNPPPELTKDDVOHKOVCFSRTSGSSEFVS
DVPNAPIPDLFPPGQITDILNAEIHGSLINLTWAPDDDHGTAKHYIIRISTIID
LRKFNESLQVNTTALIPKEANSEVEFLFKENITTFENGIDLFAIOAVKVDKSEI
SNIRVSLFIPQTPPTPPSPDETSAPOPNHINSTIPGIHILKIMKMWIGELQLSIA"
6667..6807
/gene="hclcal"
/number=3
6979..7126
/gene="hclcal"
/number=4
6966..9801
/gene="hclcal"
/number=5
15454..15631
/gene="hclcal"
/number=6
18591..18809
/gene="hclcal"
/number=7
19783..20010
/gene="hclcal"
/number=8
22253..22427
/gene="hclcal"
/number=9
24522..24628
/gene="hclcal"
/number=10

```


exon 26644..26859
/gene="hclca1"
/number=11
exon 27447..27708
/gene="hclca1"
/number=12
exon 28766..28936
/gene="hclca1"
/number=13
exon 31840..32079
/gene="hclca1"
/number=14
exon 32919..33522
/gene="hclca1"
/number=15
BASE COUNT 11245 a 6870 c 7230 g 9933 t
ORIGIN

Query Match 94.4%; Score 228.4; DB 9; Length 35278;
Best Local Similarity 99.2%; Pred. No. 5.9e-41;
Matches 240; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GTTATTCCTCCACAGACTCGCCAGAGACACCTAGTCTGATGAAAGCTGCTGCTTG 60
DB 33157 GTTATTCCTCCACAGACTCGCCAGAGACACCTAGTCTGATGAAAGCTGCTGCTTG 33216
QY 61 TCCTAATATTCATATCAACAGACACACATCTGCGATTCACATTTTAAATATATGAGAA 120
DB 33217 TCCTAATATTCATATCAACAGACACACATCTGCGATTCACATTTTAAATATATGAGAA 33276
QY 121 GTGATAGAGAACTGAGCTGTCATAGCTAGGGCTGAATTTTGTCAATATAATAA 180
DB 33277 GTGATAGAGAACTGAGCTGTCATAGCTAGGGCTGAATTTTGTCAATATAATAA 33336
QY 181 ATAAATCATTCATCCCTTTTGTGATTAATAATTTTCTAATATGATTTAGACTTCT 240
DB 33337 ATAAATCATTCATCCCTTTTGTGATTAATAATTTTCTAATATGATTTAGACTTCT 33395
QY 241 GT 242
DB 33396 GT 33397

RESULT 6
AF127036 2826 bp mRNA PRI 10-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)
DEFINITION mRNA, complete cds.
ACCESSION AF127036
VERSION AF127036.1 GI:4585468
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2826)
AUTHORS Agnel,M., Vernet,T. and Culouscou,J.M.
TITLE Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea
JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792
REFERENCE 2 (bases 1 to 2826)
AUTHORS Agnel,M. and Culouscou,J.-M.
TITLE Direct Submision
JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrieres, Rueil-Malmaison 92500, France
FEATURES
source
1..2826
/organism="Homo sapiens"
/db_xref="taxon:9606"

gene
CDs
/tissue_type="small intestine; colon"
1..2826
/gene="CaCC1"
5..2749
/gene="CaCC1"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 1"
/protein_id="A025469.1"
/db_xref="GI:4585469"
/translation="MGPKSSVPIILHILEGALSLSLIOLNNGVEIVATDPNP
EDFTLLOQIKDNTQSLVFEATGKRFRTKRVNAILIPETWKIKADYVRPKETYNNA
DVLVAESTPGNDPEYEQNGEGERHILHLPDPAGKRLAEGVPOGAFAEHNAH
LKWGEFDEYNDEKFLYSNGLQAVSICSGITQNVNKKCGSGCYKTRCTENKRYGL
YKGEFDEYLOSROTERKASIMFAOVSISYIEFCOEONHNKAPAKONKCLRSTWVI
ROSEDERKTPMTPTPPNPPEFSILOIGORIVCLIDKSGSMATGNLNRINOGOLF
LQVELEGSWGWTTPPSAAHVSSELIQINSGRDPLAKRLPAAAGGCTISCGLSLA
FTVIRKRPIDGSEIVLIDGEDNTISGCEVEKQSGALIHVALGPSAQLLEELSK
MTGILQTVASDVOVNGNLIDAFGLSSGNVAVSQRSLQESKGLTONSOMNGTVIV
DSVGRDTEFLITWTQPPQIILMDPSGKGGFVVDKNTKMAVLOIPGIAKGTWKY
SLOASOTLIVTSRASNAATLPPIVTSKNTDTSKPSPLVYVANIROGASPLRA
SVTALLESVNGKTVTEILDNGCADATDDGVYSRYFTTYDNGRYSVVRALGVN
AARRVITPOSSGALITPGATENDEIOMNPRPREINKDDVOHKVCSRSRSGSPAS
DVNPATPDLFPPGQITDLEAETHGSLNLWTAPGDDYDHTAHKYLIRISTILD
LRKFENESLOVNTTALIPKEANSEVFLRPENITFENGTDLIAIOAVDKVLSKEI
SNIAVSLFIPQTPPETPSPDETSAPCPHINISTYIPGIHLIKIMKWIGELQSLIA

BASE COUNT 875 a 623 c 632 g 696 t
ORIGIN

Query Match 88.2%; Score 213.4; DB 9; Length 2826;
Best Local Similarity 97.3%; Pred. No. 1.1e-37;
Matches 217; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GTTATTCCTCCACAGACTCGCCAGAGACACCTAGTCTGATGAAAGCTGCTGCTTG 60
DB 2596 GTTATTCCTCCACAGACTCGCCAGAGACACCTAGTCTGATGAAAGCTGCTGCTTG 2655
QY 51 TCCTAATATTCATATCAACAGACACACATCTGCGATTCACATTTTAAATATATGAGAA 120
DB 2656 TCCTAATATTCATATCAACAGACACACATCTGCGATTCACATTTTAAATATATGAGAA 2715
QY 121 GTGATAGAGAACTGAGCTGTCATAGCTAGGGCTGAATTTTGTCAATATAATAA 180
DB 2716 GTGATAGAGAACTGAGCTGTCATAGCTAGGGCTGAATTTTGTCAATATAATAA 2775
QY 181 ATAAATCATTCATCCCTTTTGTGATTAATAATTTTCTAATATGATTTAGACTTCT 223
DB 2776 ATAAATCATTCATCCCTTTTGTGATTAATAATTTTCTAATATGATTTAGACTTCT 2818

RESULT 7
AK024970 2022 bp mRNA PRI 29-SEP-2000
LOCUS Homo sapiens cDNA: FLJ21317 fls, clone COL02275, highly similar to
DEFINITION AF127036 Homo sapiens calcium-activated chloride channel protein 1
(CaCC1) mRNA.
ACCESSION AK024970
VERSION AK024970.1 GI:10437397
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_1lb:COL clone:COL02275.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NBD0 human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 2022)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,

TITLE Shihabara,T., Tanaka,T. and Nakamura,Y.
 JOURNAL Direct Submission
 Submitted (29-AUG-2000) to the DDBJ/EMBL/Genbank databases. Sunto
 Sugano, Institute of Medical Science, University of Tokyo,
 Laboratory of Genome Structure Analysis, Human Genome Center;
 Shirokane-da, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan
 (E-mail: dnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction, 5'-6 3'-end pass sequencing: Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).

FEATURES
 source
 1..2022
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COL02275"
 /clone_11b="COL"
 /tissue="colon"
 /note="Cloning vector pME18SFL3"

misc-feature
 1..2022
 /note="highly similar to AF127036 Homo sapiens
 calcium-activated chloride channel protein 1 (CACCL1) mRNA"

BASE COUNT 612 a 472 c 453 g 485 t
 ORIGIN

Query Match 81.4%; Score 197; DB 9; Length 2022;
 Best Local Similarity 100.0%; Pred. No. 4.6e-34;
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTTATTCCTCCACAGACCTCCGCCAGACACCTAGTCTGATGAAGCTGCTCTTGG 60
 |||||||
 DB 1805 GTTATTCCTCCACAGACCTCCGCCAGACACCTAGTCTGATGAAGCTGCTCTTGG 1864
 |||||||
 OY 61 TCGTATATTCATATTCATCAAGACACCTTCCTGCGATTCACATTTTAAATATGTGAA 120
 |||||||
 DB 1865 TCGTATATTCATATTCATCAAGACACCTTCCTGCGATTCACATTTTAAATATGTGAA 1924
 |||||||
 OY 121 GTGATAGAGAACTGACCTGTCAATACCTTGAGGCTGAATTTTTCAGATTAATAA 180
 |||||||
 DB 1925 GTGATAGAGAACTGACCTGTCAATACCTTGAGGCTGAATTTTTCAGATTAATAA 1984
 |||||||
 OY 181 ATTAATCATTCATCTT 197
 |||||||
 DB 1985 ATTAATCATTCATCTT 2001
 |||||||

RESULT 8
 195746
 LOCUS 195746 878 bp DNA PAT 01-DEC-1998
 DEFINITION Sequence 8 from patent US 5733748.
 ACCESSION 195746
 VERSION 195746.1 GI:3940216
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 878)
 AUTHORS Yu,G. and Rosen,C.
 TITLE Colon specific genes and proteins
 JOURNAL Patent: US 5733748-A 8-31-MAR-1998;
 FEATURES
 source
 1..878
 /organism="unknown"

BASE COUNT 257 a 179 c 188 g 241 t 13 others
 ORIGIN

Query Match 68.9%; Score 166.8; DB 6; Length 878;

Best Local Similarity 91.6%; Pred. No. 2.1e-27;
 Matches 196; Conservative 1; Mismatches 15; Indels 2; Gaps 2;

OY 1 GTTATTCCTCCACAGACTCCGCCAGACACCTAGTCTGATGAAGCTGCTCTTGG 60
 |||||||
 DB 625 GTTATTCCTCCACAGACTCCGCCAGACACCTAGTCTGATGAAGCTGCTCTTGG 684
 |||||||
 OY 61 TCGTATATTCATATTCATCAAGACACCTTCCTGCGATTCACATTTTAAATATGTGGA 119
 |||||||
 DB 685 TCGTATATTCATATTCATCAAGACACCTTCCTGCGATTCACATTTTAAATATGTGGA 744
 |||||||
 OY 120 ACTGATAGAGAACTGACCTGTCAATACCTTGAGGCTGAATTTTTCAGATTAATAA 179
 |||||||
 DB 745 ACTGATAGAGAACTGACCTGTCAATACCTTGAGGCTGAATTTTTCAGATTAATAA 803
 |||||||
 OY 180 ATTAATCATTCATCTT 213
 |||||||
 DB 804 ATTAATCATTCATCTT 837
 |||||||

RESULT 9
 AF095584 3079 bp mRNA MAN 13-OCT-2000
 LOCUS AF095584
 DEFINITION Sus scrofa epithelial chloride channel protein (AECCL1) mRNA,
 complete cds.
 ACCESSION AF095584
 VERSION AF095584.1 GI:6002645
 KEYWORDS
 SOURCE Pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 3079)
 Gasparr,K.J., Racette,K.J., Gordon,J.R., Loewen,M.E. and
 Forsyth,G.W.
 Cloning a chloride conductance mediator from the apical membrane of
 porcine ileal enterocytes
 Physiol. Genomics (Online) 3 (2), 101-111 (2000)
 20473747
 11015605
 MEDLINE
 PUBLISHED
 2 (bases 1 to 3079)
 Gasparr,K.J., Gabriel,S.E., Racette,K.J. and Forsyth,G.W.
 Direct Submission
 Submitted (28-SEP-1998) Veterinary Physiological Sciences,
 University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4,
 Canada

FEATURES
 source
 1..3079
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /tissue="ileal mucosa"
 1..3079
 /gene="AECCL1"
 133..2886
 /gene="AECCL1"
 /function="Involved in exocrine chloride conductance"
 /product="epithelial chloride channel protein"
 /protein_id="AF00077.1"
 /db_xref="GI:6002645"

CDs
 gene
 1..3079
 /gene="AECCL1"
 133..2886
 /gene="AECCL1"
 /function="Involved in exocrine chloride conductance"
 /product="epithelial chloride channel protein"
 /protein_id="AF00077.1"
 /db_xref="GI:6002645"
 /translation="MGSRSSSLFVILHLGSAQNSLIQLNGKYEIVIAIDPNP
 EDERLQIKDMVTKASPLYLEATEKREYFNVAILIPASRKAREYKPLEYKNA
 DVVVEPNPNDGPTVDMONGCEKGIPTDPVAGKRVLYQGVGFVHEHMA
 LRMGVFENYNEOKYLSNKKKQEPVIGSAAIRGNVLPCCGGSCVTRPCRADVTGL
 FQKECEFLPDPOQSRKASIMPEASIDVVERCKEKNHNRKAPNPNONOKNLSRMEVI
 ODSEPKTTPTPTTPPTPTSLDIOGRVICYLDKSGSTVGRKRLQDAKTEL
 LQVBOGWAVAVDAIVKSELVQINSARALAKSLPTASGETSICSGLRSA
 FTVIRKKYPTDSESLVLTLDDEDNTISCFEVRONGAIHTVALGPSAAELELSO
 MTGGLQTVASQDAENNGLLIDAFGLSSGNRASORSIOLESQGLTLQNNEMNGTVV
 DSTVQKDLFLITLERKFLSPILPFGVPSGRSODSLVGHKHKMAVEQVPGTAKVMV
 KYSLOASSOTLTLVSSRRSSATLPVPTVTSKMNKDTGKPPSPVAVYKIHQGLPLIL
 RAKVTALIESENGKVTLELLDNGAGATATNDGITYSRFTAYANGRYSKVAVNLGG
 VNTPRRRAPPLMSGAMTIRGNIESEIKMNPDPIDKDLQAGVCSRTASGSPV

ASDVKSPSPDLPPCKITLTDKAGIOLNLTWPAEDDYDGHGRADRIIRISINI
LDLRKFNDSVQVNTTDLIPKANSBEVFEVPEEIPINGDLPLAVOAVDTINKS
EISNTAQVSLFPEAPPEPETPAPSLPCEIIVNSTIPIGHIILKIMKWLGEIOL
SIA"

BASE COUNT 892 a 758 c 755 g 674 t

Query Match 52.0%; Score 125.8; DB 4; Length 3079;
Best Local Similarity 77.9%; Pred. No. 2.5e-18;
Matches 183; Conservative 0; Mismatches 37; Indels 15; Gaps 2;

OY 7 TCCTCCACAGACTCCGCCAGACACCTTACCTGTATGAAAGCTGCTCTCTCTCTAA 66
DB 2748 TCCTCCGAGACTCTCTCGGAGACCTGCTCTCTCTCTCTCTCTCTCTCTCTGA 2798
OY 67 TATTCATACACAGACCATTCCTGCTATTCATTTTAAATATGGAAGTGGAT 126
DB 2799 AATTCAGTCACAGACCATTCCTGCTATTCATTTTAAATATGGAAGTGGAT 2858
OY 127 AGGAGAACTGACAGCTGCTGATAGCTGATGATTTTGTGATTAATAATAAT 186
DB 2859 GGGAGAAATACAGCTATTCATTCAGCTAGGCTGATTTTCTGATTAATAATAATCA 2915
OY 187 CATTCATCTTTTGTATTAATAATTTCTTAATATGATTTAGACTCTCTG 241
DB 2916 ---TCCATCTCTGCTTTGATTAATAATAATTTAGACTCTCTG 2967

RESULT 10
AB017156 2937 bp mRNA ROD 10-NOV-1999

LOCUS AB017156
DEFINITION Mus musculus gob-5 mRNA, complete cds.
ACCESSION AB017156
VERSION AB017156.1 GI:3721911
KEYWORDS GOB-5.
SOURCE Mus musculus adult intestine goblet cell cDNA to mRNA.
ORGANISM Mus musculus

REFERENCE 1 (sites)
AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.
TITLE Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice
JOURNAL Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)
99160866

2 (bases 1 to 2937)
REFERENCE 2
AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1998) to the DDBJ/EMBL/Genbank databases. Tohru Komiyama, ERATO, JST, Genome Asymmetry Group, Doi Biosymmetry Project, Tokodai 5-9-9, Tsukuba, Ibaragi 300-2635, Japan (E-mail: tkomied@erauto.ttc-net.co.jp, Tel: 81-298-48-1515, Fax: 81-298-47-8901)

FEATURES
source location/Qualifiers

1. .2937
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="goblet cell"
/dev_stage="adult"
/tissue_type="intestine"
15. .2756
/gene="gob-5"
15. .2756
/gene="gob-5"
/codon_start=1
/protein_id="BA33743.1"
/db_xref="GI:3721912"
/translation="MESLSPVFLILHLLEGLVSLIOLNNNGYEGIVAIIDHVP
EDALIOHKVTOASPLSPKATGKRFYKALIPESMKALIPESKAPETPKLLETNA
DVLVSTSPGLNDEPTEHIGAGKGIHLTPPLAGKTLTGVGPODRVTHENAH
FRMGVFNENDEKRYLSEKQAVRCSALITGKNVRRGCGSCTTNKRCYIDRTVEY
LTKDNCFVPPDHONEKASIMFNQINISVETFEKHNQNEADPNQNCNLRSTWEY

BASE COUNT 860 a 718 c 693 g 666 t

Query Match 38.0%; Score 92; DB 10; Length 2937;
Best Local Similarity 67.1%; Pred. No. 7.2e-11;
Matches 149; Conservative 0; Mismatches 65; Indels 8; Gaps 1;

OY 21 CGCCAGACACCTAGTCTGATGAAACGTCTCTCTCTCTCTCTCTCTCTCTCTCA 80
DB 2617 CTCAGAGCGCCCATTCGCCGAGACTCACTCCCTGTCTGACATCAGCATCA 2676
OY 31 GCACATTCCTGCTACATTTTAAATATGGAAGTGGATGAGAACTGAC 140
DB 2677 GCACATTCCTGCTACATTCATTTTAAATATGGAAGTGGATGAGAACTGAC 2736
OY 141 TGTCAATGCTGAGGCTGATTTTGTGATTAATAATAATATCAATTCATCTCT 200
DB 2737 TGACACTAGGTTGACATGATTTTACGCAACAACAACA-----GTCAATTCCT 2788
OY 291 TTGATTTTAAATTTCTTAAATGATTTTATGACTCTCTCT 242
DB 2789 TTCACTGGAAATTTCTTAAATGATTTTATGACTCTCTCT 2830

RESULT 11
AC005139 256172 bp DNA HTG 01-APR-1999
LOCUS AC005139
DEFINITION Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS
ACCESSION AC005139
KEYWORDS HTG; HTGS; PHASE1.
SOURCE malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum

REFERENCE 1 (bases 1 to 256172)
AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurd, O.B. and Davis, R.W.
TITLE Plasmodium falciparum 3D7 chromosome 12
JOURNAL Unpublished
2 (bases 1 to 256172)
REFERENCE 2
AUTHORS Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

On Apr 2, 1999 this sequence version replaced gi:4337170.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 126535: contig of 126535 bp in length
* 126536 126735: gap of unknown length
* 126736 158845: contig of 32110 bp in length
* 158846 159045: gap of unknown length
* 159046 211538: contig of 52493 bp in length
* 211539 211738: gap of unknown length
* 211739 248106: contig of 36368 bp in length


```
misc_feature /chromosome="13"
repeat_region /clone="RP11-522N4"
misc_feature /clone_11b="RC1-11.2"
1. .497
/note="match: GSS: Em:A0591995"
1249. .1328
repeat_region /note="40 copies 2 mer tt 66% conserved"
misc_feature complement(1368. .1575)
repeat_region /note="match: STS: Em:G11815"
1421. .1782
/note="MER21B repeat: matches 14. .375 of consensus"
1783. .2038
repeat_region /note="MER21B repeat: matches 506. .794 of consensus"
2033. .2112
/note="MER21B repeat: matches 703. .775 of consensus"
2114. .2155
repeat_region /note="21 copies 2 mer at 76% conserved"
4505. .5520
repeat_region /note="L1MA10 repeat: matches 5370. .6308 of consensus"
5529. .5843
misc_feature /note="match: STS: Em:G31647"
5844. .6408
repeat_region /note="L1MA10 repeat: matches 4785. .5370 of consensus"
6187. .8278
repeat_region /note="46 copies 2 mer aa 63% conserved"
10792. .10881
repeat_region /note="15 copies 6 mer gaaaga 71% conserved"
10829. .10892
repeat_region /note="32 copies 2 mer ag 73% conserved"
10921. .10974
repeat_region /note="27 copies 2 mer gt 77% conserved"
10923. .10976
repeat_region /note="9 copies 6 mer gttgtg 75% conserved"
12312. .12410
repeat_region /note="L1MA4 repeat: matches 5336. .5435 of consensus"
12395. .13218
repeat_region /note="L1MA4 repeat: matches 5414. .6263 of consensus"
15363. .15431
repeat_region /note="L1MA4 repeat: matches 6206. .6274 of consensus"
18369. .18434
repeat_region /note="11 copies 6 mer acacac 74% conserved"
18375. .18434
repeat_region /note="30 copies 2 mer ac 76% conserved"
25426. .25485
repeat_region /note="10 copies 6 mer caaaa 71% conserved"
25436. .25513
repeat_region /note="39 copies 2 mer aa 67% conserved"
25704. .25749
repeat_region /note="23 copies 2 mer ac 84% conserved"
25708. .25749
repeat_region /note="7 copies 6 mer acacac 88% conserved"
complement(27993. .28572)
misc_feature /note="match: GSS: Em:A0742318"
complement(28021. .28568)
misc_feature /note="match: GSS: Em:AQ749344"
complement(28021. .28560)
misc_feature /note="match: GSS: Em:A0545310"
complement(28031. .28613)
misc_feature /note="match: GSS: Em:AQ471460"
28263. .28276
repeat_region /note="MER50 repeat: matches 289. .733 of consensus"
29927. .30183
repeat_region /note="LTR9B repeat: matches 306. .561 of consensus"
30480. .30781
/note="LTR9B repeat: matches 1. .306 of consensus"
30783. .33414
repeat_region /note="HUES-R-P3b repeat: matches 5105. .7411 of consensus"
31572. .32570
/note="MER61-internal repeat: matches 2338. .3304 of consensus"
32875. .33701
repeat_region /note="MER61-internal repeat: matches 1502. .2338 of consensus"
33701. .33798
repeat_region /note="HUES-R-P3b repeat: matches 3796. .3874 of consensus"
34106. .34174
repeat_region /note="HUES-R-P3b repeat: matches 3722. .3796 of consensus"
34178. .34706
repeat_region /note="HUES-R-P3b repeat: matches 3073. .3591 of consensus"
34707. .35097
repeat_region /note="LTR8 repeat: matches 1. .402 of consensus"
35098. .35188
repeat_region /note="HUES-R-P3b repeat: matches 2565. .2655 of consensus"
35186. .35644
repeat_region /note="HUES-R-P3b repeat: matches 747. .2292 of consensus"
36591. .36735
repeat_region /note="HUES-R-P3b repeat: matches 409. .559 of consensus"
36739. .36794
repeat_region /note="HUES-R-P3b repeat: matches 46. .96 of consensus"
36815. .36834
repeat_region /note="MER61B repeat: matches 544. .563 of consensus"
36835. .37034
repeat_region /note="LTR9B repeat: matches 445. .644 of consensus"
37593. .38028
repeat_region /note="Tiger2a repeat: matches 5. .434 of consensus"
38828. .39878
repeat_region /note="MER11A repeat: matches 1. .1266 of consensus"
41787. .41864
repeat_region /note="39 copies 2 mer at 65% conserved"
41989. .42052
repeat_region /note="32 copies 2 mer aa 68% conserved"
43089. .43124
repeat_region /note="18 copies 2 mer ac 88% conserved"
43544. .44276
repeat_region /note="LTR8 repeat: matches 1. .691 of consensus"
50003. .51265
repeat_region /note="L1P repeat: matches 3220. .4490 of consensus"
51269. .51497
repeat_region /note="L1PA15 repeat: matches 5918. .6155 of consensus"
51806. .52730
repeat_region /note="L1MA7 repeat: matches 5336. .6287 of consensus"
complement(52848. .53327)
misc_feature /note="match: GSS: Em:AQ883194"
52908. .53407
repeat_region /note="MER67D repeat: matches 1. .466 of consensus"
53638. .54020
repeat_region /note="MER57B repeat: matches 16. .403 of consensus"
56734. .57210
repeat_region /note="MER31B repeat: matches 1. .481 of consensus"
complement(58958. .59437)
misc_feature /note="match: GSS: Em:AQ772237"
59374. .59792
repeat_region /note="LTR26 repeat: matches 1. .430 of consensus"
59793. .60150
repeat_region /note="THEIC repeat: matches 1. .371 of consensus"
60154. .61563
repeat_region /note="THEIC-internal repeat: matches 3. .1577 of consensus"
61566. .61942
repeat_region /note="THEIC repeat: matches 3. .371 of consensus"
61948. .62105
repeat_region /note="LTR26 repeat: matches 427. .601 of consensus"
63298. .63678
repeat_region /note="THEIC repeat: matches 1. .371 of consensus"
63684. .65289
repeat_region /note="THEIC-internal repeat: matches 4. .1575 of consensus"
65297. .65643
repeat_region /note="THEIC repeat: matches 1. .367 of consensus"
complement(65870. .66492)
misc_feature /note="match: GSS: Em:AQ002365"
73059. .75072
repeat_region /note="L1PA3 repeat: matches 4139. .6146 of consensus"
75074. .75656
repeat_region /note="L1PA3 repeat: matches 3561. .4148 of consensus"
complement(76808. .77333)
```


[illegible]

intron complement(14298..14380)
/number=2
exon complement(14381..14461)
/number=3
intron complement(14462..14651)
/number=3
exon complement(14652..14731)
/number=4
intron complement(14732..14824)
/number=4
exon complement(14825..14918)
/number=5
intron complement(14919..15039)
/number=5
exon complement(15040..15131)
/number=6
intron complement(15132..15317)
/number=6
exon complement(15318..15408)
/number=7
intron complement(15409..15493)
/number=7
exon complement(15494..15549)
/number=8
gene /number=8
15547..17851
/gene="r8B10_30"
complement(join(16547..16646,17116..17228,17358..17467,
17581..17851))
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB81823.1"
/db_xref="GI:7287985"
translation="MVKPAFLIDNPFSPFEDSGDKAMSFNCSLSRRSLVYVIVASPC
LLPALSSAKTKSPYDERLLRNFONRIBENNAPDEPFVREGGPKVLAADNY
IKADSGLYRDPNCGQPPKDCQVTHYICNSGRIDSTYIQGSPARIMGTNA
LVPGRSKQFEVDVDELSTIONCERTIIGFYSVYCS"
complement(16547..16646)
/number=1
intron complement(16647..17115)
/number=1
exon complement(17116..17228)
/number=2
intron complement(17229..17357)
/number=2
exon complement(17358..17467)
/number=3
intron complement(17468..17580)
/number=3
exon complement(17581..17851)
/number=4
gene /number=4
18460..20691
/gene="r8B10_40"
complement(18460..20691)
/note="contains ATP-dependent DNA ligase signatures
AA577-585"
/codon_start=1
/product="putative protein"
/protein_id="CAB81824.1"
/db_xref="GI:7287986"
translation="MASPNPTKRRSPNVVPPPRKSTIGGCFECKVLPALFL
ALPLPSQADPVEGTVLTKFMELIHLFVGLAVYGLFSRRNVESAVDLNTRVDES
SLSYVRIEYQVSVDEEDNDSEFVVRSDSVASAVYKSESFEVSEGELES
SEFETNEVRAMNSQYFOGKSGVYVAVGADGVHVOPLGLPIRRLSSSLADNAAL
ODKSPADSCDGVNAEASLADNPFDEVLAPADPVMQARPEMNGIGDNPSPROP
ISVDPTLKSISRSTGSSSSQTSYASQONRSPRSYSAELSNVBEILVYKESROS
SSRSSSPPLPSPLSPSPPELVPNDRRSPPLVTDPTFRASHSHYSDGSLLE
EDVRGFEENELGSKVGRKAEFFSKERKSGSLAAESSRRGRKSRSSYPSEISS
PVGAGDSTTRQDLOQKSNCHLEENIRKGVADHNNLRVYKGRSHSLTJADSA
KDEKYESFPALDVYOPTNAASRRARSSRGSDTLPEKDVYRKLDDVDSDKS
RKDLPGKDKEMKLDGPRLEPSRRASSVSLRGKSVRTIRSDRGRKDKVTKDGSDS
RAKAEVSRGRTRKSRPROEELSTVHOKESSEMPAKGPEVAMEEPPQAEQDPVTF
EEEEAAMESSQSNASHDINEYDKRAGEPTAKREQIRIQKLISGQPPRGGCTGIRNS
QFR"

exon complement(18460..20691)
/number=1
gene 22335..23917
/gene="r8B10_50"
complement(join(22335..22644,23038..23117,23202..23560,
23719..23917))
/note="Contains 'Homeobox' domain signature and profile
AA192-215"
/codon_start=1

Query Match 17.3%; Score 41.8; DB 8; Length 103787;
Best Local Similarity 54.1%; Pred. No. 10;
Matches 85; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 137 CACCTGTCATACCTAGGCTGAATTTTTCAGATAAATAAATTAATCATTCATCCT 196
DB 26603 AACATTTCTTACCTACCTATCTTATATATCAAGTAGAAGTAGAGAAATTT 26662
QY 137 CACCTGTCATACCTAGGCTGAATTTTTCAGATAAATAAATTAATCATTCATCCT 196
DB 26603 CAAATATCATCATATATAGCTTACCTTTTTCATATGAACTACAAACTCAACACTTTT 26722
QY 197 TTTTTCGATTATATAAATTTTCTAAATGATTTTTCAG 233
DB 26723 TTTTTCGTTGAAAGAAATTCAAAGTTTCATATTTTG 26759

RESULT: 115
AF214670 678 bp DNA INV 23-MAY-2001
LOCUS Apis mellifera cytochrome oxidase II (COII) gene, complete cds;
DEFINITION mitochondrial gene for mitochondrial product.
ACCESSION AF214670
VERSION AF214670.1 GI:11990602
KEYWORDS honeybee.
SOURCE Mitochondrion Apis mellifera
ORGANISM Eukaryota; Metazoa; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
Aculeata; Apoidea; Apidae; Apis.

REFERENCE 1 (bases 1 to 678)
Tanaka, H., Roubik, D.W., Kato, M., Llew, F. and Gunsalam, G.
AUTHORS phylogenetic position of Apis nulensis of northern Borneo and
TITLE phylogeography of A. cerana as inferred from mitochondrial DNA
sequences
JOURNAL Insects Soc. 48, 44-51 (2001)
REFERENCE 2 (bases 1 to 678)
Tanaka, H., Kato, M. and Roubik, D.W.
AUTHORS Direct Submission
TITLE Submitted (13-DEC-1999) Center for Ecological Research, Kyoto
JOURNAL University, Kitashirakawa Nishimachi, Kyoto 606-8502, Japan

FEATURES
source Location/Qualifiers

gene 1..678
/organism="Apis mellifera"
/organelle="mitochondrion"
/db_xref="taxon:7460"
1..678
/gene="COII"

CDS 1..678
/gene="COII"
/codon_start=1
/transl_table=5
/product="cytochrome oxidase II"
/protein_id="AAC42211.1"
/db_xref="GI:11990603"

translation="MSTWFMFQESNSYVADNLISFHMVMIIIMISTLVYIILD
LPMNFSNLFKNHIEIITWIIPIIILICFSLKILYIDELVNFESKISGH
QWYASYEYEFNNIEDSYMLVNNINCRLETFNRMYIPKPIPLATITSDVYIS
HWVPSIGIVDAVGRINQNLNISKPGIFFQGCSEICMNNISFMP IWEVNSFVFL
MNVNVOI"

BASE COUNT 265 a 77 c 56 g 280 t

Thu Apr 4 09:27:32 2002

us-09-049-696-16.rge

Page 12

Query Match	16.9%	Score 41;	DB 3;	Length 678;
Best Local Similarity	52.7%	Pred. No. 12;		
Matches	89;	Conservative	0;	Mismatches 80;
				Indels 0;
				Gaps 0;

Oy	61	TCTTAATATTCATATTCACACAGCCACCATCTCTGCGATTCACAAATTTTAAATTTGTGGGAA	120
	111		
	174	TCATTAATTTGTAATAATTTATTTTAAACAATTAATTCACATTAATTAATTAATTAATTTGGTT	223
	174		
Oy	121	GTGGATTAGGAGAACTCCACGCTGCATAAGCCTAGCGCTGCAATTTTGTCCAGATAAATAAA	180
	121		
Db	234	TCCATCATCTTAAAAATTTTATTTTAAATTTGATGAAATTTGTAAATCCTTTTTTTTCAAATTA	293
	234		
Oy	181	ATTAATTCATTCATCCCTTTTTTTTGATTTATAAAATTTTCTAAATGTATT	229
	181		
Db	234	ATCAATTTGGTCATCAATCATATTTGATCTCATTAAGCAATATCCAGAAATTTTAAT	342
	234		

Search completed: April 4, 2002, 03:07:59
Job time: 60761 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:43 ; Search time 16681 Seconds
(Without alignments)
99.205 Million cell updates/sec

Title: US-09-049-696-15

Perfect score: 154
Sequence: 1 ACAGATCTTTCATTGCTAT.....TCCTTGCTCAATATTCATA 154

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_est1:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_tod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154	100.0	335	10	AI721121 as73d08.x
2	154	100.0	378	10	AA296955 EST112726
3	154	100.0	490	10	AA508854 n122h10.s
4	154	100.0	501	10	AI721275 as82h08.x
5	154	100.0	504	10	AI802693 w17c01.x
6	154	100.0	712	10	AM009763 ws87b04.x
7	154	100.0	716	10	AI660234 w68g02.x
8	154	100.0	734	10	AI660957 w12d008.x
9	152.4	99.0	561	10	AI602756 w18a04.x
10	151	98.1	450	10	AM009764 ws87b05.x
11	150.8	97.9	466	11	BG195179 RST14338
12	112	72.7	255	11	D25727 HDNGS04094

C 13	90.8	59.0	392	10	AI687981 wa77f06.x
C 14	75.4	49.0	405	11	BI292301 UI-R-DNO-
C 15	75.4	49.0	435	11	BE995128 UI-M-CGP
C 16	75.4	49.0	469	11	BI301140 UI-R-DNO-
C 17	75.4	49.0	475	11	BG372819 UI-R-CVI-
C 18	75.4	49.0	512	11	BI291875 UI-R-DNO-
C 19	75.4	49.0	555	11	BI291875 UI-R-DNO-
C 20	75.2	48.8	175	10	AI506262 vq50e03.x
C 21	73.6	47.8	297	10	AI641931 vq50e03.y
C 22	73.6	47.8	301	10	AA689856 vt62c11.r
C 23	73.6	47.8	388	10	AA734161 vs19g06.r
C 24	73.6	47.8	418	10	AA596289 vo2a08.r
C 25	73.6	47.8	421	10	AA691586 vs12g06.r
C 26	73.6	47.8	431	10	AA338284 mv34g09.r
C 27	73.6	47.8	463	10	AA839323 vo40f07.r
C 28	73.6	47.8	482	10	AA711228 vt70c11.r
C 29	73.6	47.8	488	10	AA733415 vt75b06.r
C 30	73.6	47.8	552	10	AA688953 vs04h07.r
C 31	73.6	47.8	646	10	AA691335 vs14h06.r
C 32	73.6	47.8	732	11	BG963719 602831354
C 33	73.6	47.8	2915	12	AK007466 Mus muscu
C 34	73.6	47.8	2933	12	AK008659 Mus muscu
C 35	72	46.8	629	10	AA518112 v123c06.r
C 36	72	46.8	812	11	BG968878 602836186
C 37	70.6	45.8	386	10	AA692521 vt59g08.r
C 38	69.8	45.3	452	10	AA822200 vp36d05.r
C 39	68.6	44.5	207	10	AA669318 vq50e03.r
C 40	68.6	44.5	507	10	AA623217 v14d07.r
C 41	68.6	44.5	554	10	AA623196 v14b07.r
C 42	65.4	42.5	310	10	AA297150 EST112734
C 43	65	42.2	258	11	BF523490 UI-R-YO-a
C 44	63	40.9	205	10	AI675991 wd08g07.x
C 45	62.8	40.8	687	11	BG286238 602383065

ALIGNMENTS

RESULT: 1
LOCUS AI721121/c
DEFINITION as73d08.x1 Barsstead colon HPLR87 Homo sapiens CDNA clone
IMAGE:2334351 3' similar to TR:088826 O88826 GOB-5 PROTEIN. ;, mRNA
sequence.

ACCESSION AI721121
VERSION AI721121.1 GI:5038377
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 335)

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,

Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin

, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,

White, Y., Wylie, T., Waterston, R. and Wilson, R.

WashU-NCI human EST Project

Unpublished (1997)

CONTACT: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from gibco

High quality sequence stop: 312.

FEATURES

source 1..335

/organism="Homo sapiens"

/db_xref="taxon:9606"


```

/clone="IMAGE:2334351"
/clone.lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/notes="Organ: Colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site:1: EcoRI; Site:2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCAGAACTGCAAGTGGAGCGCCGCTTTTGTATTCCTCCACAGACTCCGACAGACCTAGTCTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[5' ATTACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested
with Not I and cloned into the Not I and Eco RI sites of
the modified pT73 vector. Library constructed by Bob
Barstead."

```

BASE COUNT 98 a 57 c 71 g 109 t

ORIGIN

Query Match 100.0%; Score 154; DB 10; Length 335;
 Best Local Similarity 100.0%; Pred. No. 9.3e-37;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ACAGATCTTTTCATTCGATTCAGGCTGTGATTAAGCTGATCGAATCAGAAATATCC 60
    |||
DB 275 ACAGATCTTTTCATTCGATTCAGGCTGTGATTAAGCTGATCGAATCAGAAATATCC 216
    |||
OY 61 AACATTGCACGAGTATCTTTGTTTATTCCTCCACAGACTCCGACAGACCTAGTCTT 120
    |||
DB 215 AACATTGCACGAGTATCTTTGTTTATTCCTCCACAGACTCCGACAGACCTAGTCTT 156
    |||
OY 121 GATGAACGTCGCTGCTTCCTGCTTAATATTCATA 154
    |||
DB 155 GATGAACGTCGCTGCTTCCTGCTTAATATTCATA 122
    |||

```

RESULT 2
 AA296955 378 bp mRNA EST 18-APR-1997
 LOCUS EST112726 Colon I Homo sapiens cDNA 5' end, mRNA sequence.
 DEFINITION AA296955
 ACCESSION AA296955.1 GI:1949515
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 378)

REFERENCE
 AUTHORS
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulder,R.A., Bult,
 C.J., Lee,N.H., Kirkness,E.F., Weinstein,K.G., Gocayne,J.D., White,
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,
 L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodde,A.,
 Guelm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.T., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pellierino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Sander,D.M., Shirley,R.,
 Small,R.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dime,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A., He,W.W.,
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
 Wel,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,
 M.R., Rosen,C.A., Haseeltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 96026280
 Other ESTs: TH0167738
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056
 Fax: 3018699423
 Email: arkerlavet@igr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hg1/hg1.html>)
 Seq primer: M13 Reverse.
 Location/Qualifiers

FEATURES

source

1. 378

/organism="Homo sapiens"

/db_xref="ATCC (host):127834"

/db_xref="taxon:9606"

/clone.lib="Colon I"

/dev_stage="adult"

/note="Organ: colon; Vector: pBluescript SK-; Site:1:
 EcoRI; Site:2: XhoI"

BASE COUNT 112 a 79 c 64 g 119 t 4 others

ORIGIN

Query Match 100.0%; Score 154; DB 10; Length 378;
 Best Local Similarity 100.0%; Pred. No. 9.5e-37;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ACAGATCTTTTCATTCGATTCAGGCTGTGATTAAGCTGATCGAATCAGAAATATCC 60
    |||
DB 34 ACAGATCTTTTCATTCGATTCAGGCTGTGATTAAGCTGATCGAATCAGAAATATCC 153
    |||
OY 61 AACATTGCACGAGTATCTTTGTTTATTCCTCCACAGACTCCGACAGACCTAGTCTT 120
    |||
DB 154 AACATTGCACGAGTATCTTTGTTTATTCCTCCACAGACTCCGACAGACCTAGTCTT 213
    |||
OY 121 GATGAACGTCGCTGCTTCCTGCTTAATATTCATA 154
    |||
DB 214 GATGAACGTCGCTGCTTCCTGCTTAATATTCATA 247
    |||

```

RESULT 3
 AA508854 490 bp mRNA EST 18-AUG-1997
 LOCUS n122h.0.s1 NCI-CGAP_Co4 Homo sapiens cDNA clone IMAGE:968803 3',
 DEFINITION AA508854
 ACCESSION AA508854
 VERSION AA508854.1 GI:2246357
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 490)

REFERENCE
 AUTHORS
 TITLE
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strussberg, Ph.D.
 Email: c98pb@remail.nih.gov

Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.,
 Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LML at:
www.bio.lml.gov/brp/image/image.html
 Insert Length: 856 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 404.
 Location/Qualifiers

1. 490
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:968803"
 /clone.lib="NCI-CGAP_Co4"


```

BASE COUNT      155 a      82 c      89 g      164 t
ORIGIN
        /sex="pooled"
        /issue_type="colon"
        /lab_host="DH10B"
        /note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from pooled colon
tumor tissue, and was then primed with a Not I - oligo(dT)
primer. Double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
This library is not normalized. Library constructed by
Bento Soares and M. Fatima Bonaldo. "

```

	Query Match	Similarity	Score 154;	DB 10;	Length 490;
	Best Local	Similarity	100.08;	Pred. No. 9.7e-37;	
	Matches 154;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ACAGATCTTTTCATTCGATTCAGGCTGTGATTAAGTCGATTCGAATAATCAGAAATATCC	60		
Db	488	ACAGATCTTTTCATTCGATTCAGGCTGTGATTAAGTCGATTCGAATAATCAGAAATATCC	429		
QY	61	AACATTCGACGAGTATCTTTTGTTAATTCCTCACAAGACTCGCCACAGACACTAGTCTCT	120		
Db	428	AACATTCGACGAGTATCTTTTGTTAATTCCTCACAAGACTCGCCACAGACACTAGTCTCT	369		
QY	121	GATGAAGCTGCTGCTTCCTTAATATTCATA	154		
Db	368	GATGAAGCTGCTGCTTCCTTAATATTCATA	335		

LOCUS	A1721275	501 bp	EST	10-JUN-1999		
DEFINITION	aa82h08.x1 Barstead colon HRPB7 Homo sapiens cDNA clone IMAGE:2335263 3' similar to SW:ECIC_BOVIN P54281 EPITHELIAL CHLORIDE CHANNEL PROTEIN ;, mRNA sequence.					
ACCESSION	A1721275					
VERSION	A1721275.1	GI:5038531				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi					
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 501)					
AUTHORS	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S., Krizman,D., Kucaba,T., Lacey,M., Le,N., Lennon,G., Marra,J., Marti,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisling,B., White,X., Wylie,T., Waterston,R. and Wilson,R.					
TITLE	WASHU-NCI human EST project					
JOURNAL	unpublished (1997)					
COMMENT	Contact: Wilson RK					

Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from g1bco
High quality sequence stop: 394.
Location/Qualifiers
1..501

```

1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2335263"
/clone_id="Barstead colon HPLuB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/notes="Origin: colon; Vector: pT773D-Pac (Pharmacia) with a

```

modified polylinker: Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTACCAATCTGAGTGGAGACGGCCGCCCTTTTTTTTTTTTTTTTTT 3'); double-stranded cDNA was ligated to Eco RI adapters (5' AATTACTACTACTAAT 3' and 5' ATTACTAGTG 3'), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."

BASE COUNT	158 a	83 c	106 g	154 t
ORIGIN				

Query Match 100.0%; Score 154; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 9, 8e-37;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match	100.0%	Score 154:	DB 10:	Length 501:
Best Local Similarity	100.0%	Pred. No. 9.8e-37:		
Matches 154:	conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY 1 ACAGATCTTTTCATTGCTATTCAGGCTGTGATTAAGTCGATCTGAATACAGAAATATCC	60			
DB 254 ACAGATCTTTTCATTGCTATTCAGGCTGTGATTAAGTCGATCTGAATACAGAAATATCC	235			
QY 51 AACATTGCACGAGATATCTTTTGTATTTCTCCACACAGACTCGGCCACAGACACTAGTCTT	120			
LS 234 AACATTGCACGAGATATCTTTTGTATTTCTCCACACAGACTCGGCCACAGACACTAGTCTT	175			
QY 121 GATGAAGCTGCTGCTCTGTCCTAATATTCATA	154			
DB 174 GATGAAGCTGCTGCTCTGTCCTAATATTCATA	141			

RESULT 5
 AL802693/c
 LOCUS
 DEFINITION
 AI802693 504 bp mRNA EST 18-DEC-1999
 w17c01.x1 Soares_Dieckgreffe.colon_NHUC Homo sapiens cDNA clone
 IMAGE:23350848 3' similar to FR:088826 088826 GOB-5 PROTEIN. ; , mRNA
 sequence.
 ACCESSION
 AI802693
 VERSION
 AI802693.1 GI:5368165
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 504)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 UNPUBLISHED (1997)
 COMMENT
 Contact: Robert Stransberg, Ph.D.
 Email: cgapbs-remail.nlm.nih.gov
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 727 Std Error: 0.00
 Seq primer: -40up from Gibco
 High quality sequence stop: 392.
 FEATURES
 CDS
 Location/Qualifiers
 1..504

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2350848"
/clone_lib="Soares_Dieckgraefe_colon_NHUC"
/tissue_type="colonic mucosa from 5 ulcerative colitis patients"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5 TGTACCAATCTGAAGTCGAGCGGCCGCCTTTTCTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieck@lm.wustl.edu); colonic mucosa represents a range of

```


disease involvement from mild cryptitis to severe ulceration, fibrosis, and degeneration. Library constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 150 a 85 c 103 g 163 t 3 others

ORIGIN

Query Match 100.0%; Score 154; DB 10; Length 504;
Best Local Similarity 100.0%; Pred. No. 9, 8e-37;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACAGATCTTTTCATTGCTATTCAGGCTGTGATTAAGTCGATCGAATATATCC 60
|||||
DB 291 ACAGATCTTTTCATTGCTATTCAGGCTGTGATTAAGTCGATCGAATATATCC 232
|||||
OY 61 AACATTGCACGAGTATCTTGTATTCCTCCACAGACTCCGCGACAGACACTAGTCT 120
|||||
DB 231 AACATTGCACGAGTATCTTGTATTCCTCCACAGACTCCGCGACAGACACTAGTCT 172
|||||
OY 121 GATGAACGCTGCTCTGCTGCTATATTCATA 154
|||||
DB 171 GATGAACGCTGCTCTGCTGCTATATTCATA 138
|||||

RESULT 6
AM009763 712 bp mRNA EST 08-MAR-2000
LOCUS ws87b04.x1 NCI-CGAP_C03 Homo sapiens cDNA clone IMAGE:2504911 3'
DEFINITION similar to TR:088826 088826 GOB-5 PROTEIN.; mRNA sequence.
ACCESSION AM009763
VERSION AM009763.1 GI:5858541
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 712)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-rt@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Insert length: 807 Std Error: 0.00
Seq primer: -400P from Glibco
High quality sequence stop: 451.
Location/Qualifiers

FEATURES
SOURCE

1. 712
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2504911"
/clone_id="NCI-CGAP_C03"
/sex="Pooled"
/tissue_type="colon"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from 12 pooled bulk tumor samples and primed
with a Not I - oligo(dT) primer. Double stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library went through one round of
normalization."

BASE COUNT 217 a 130 c 143 g 222 t

Query Match 100.0%; Score 154; DB 10; Length 712;
Best Local Similarity 100.0%; Pred. No. 1e-36;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACAGATCTTTTCATTGCTATTCAGGCTGTGATTAAGTCGATCGAATATATCC 60
|||||
DB 448 ACAGATCTTTTCATTGCTATTCAGGCTGTGATTAAGTCGATCGAATATATCC 389
|||||
OY 61 AACATTGCACGAGTATCTTGTATTCCTCCACAGACTCCGCGACAGACACTAGTCT 120
|||||
DB 398 AACATTGCACGAGTATCTTGTATTCCTCCACAGACTCCGCGACAGACACTAGTCT 329
|||||
OY 121 GATGAACGCTGCTCTGCTGCTATATTCATA 154
|||||
DB 328 GATGAACGCTGCTCTGCTGCTATATTCATA 295
|||||

RESULT 7
A1660234/c 716 bp mRNA EST 18-DEC-1999
LOCUS ws68g02.x1 Soares-Dieckgraefe_colon NHCD Homo sapiens cDNA clone
DEFINITION IMAGE:2346290 3' similar to TR:088826 088826 GOB-5 PROTEIN.; mRNA
sequence.
ACCESSION A1660234
VERSION A1660234.1 GI:4763804
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 716)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-rt@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 956 Std Error: 0.00
Seq primer: -400P from Glibco
High quality sequence stop: 387.
Location/Qualifiers

FEATURES
SOURCE

1. 716
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2346290"
/clone_id="Soares-Dieckgraefe_colon_NHCD"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCCATCTGAGTGGGCGCGCTTTTATTTTATTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraefe (Washington University,
dieckel@wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Soares and M. Fatima Bonaldo.

BASE COUNT 220 a 123 c 140 g 233 t

Query Match 100.0%; Score 154; DB 10; Length 716;
Best Local Similarity 100.0%; Pred. No. 1e-36;

	MATCHES	154:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
OY	1	ACAGATCTTTTCATTGCGTATTCAGCGCTGTGGTAAGAAGTCGATCGAATAATCAGAAATATCC	60							
DB	485	ACAGATCTTTTCATTGCGTATTCAGCGCTGTGGTAAGAAGTCGATCGAATAATCAGAAATATCC	426							
OY	61	AACATTGCACGAGTATCTTTGTTATTCCTCCACAGACTCGGCCGACACACCTGTGCT	120							
DB	425	AACATTGCACGAGTATCTTTGTTATTCCTCCACAGACTCGGCCGACACACCTGTGCT	366							
OY	121	GATGAAGCTGCTGCTCTTGCTCTCAATATTCATA	154							
DB	365	GATGAAGCTGCTGCTCTTGCTCTCAATATTCATA	332							
RESULT	8									
LOCUS	A1660957/c									
DEFINITION	A1660957	734 bp	mRNA	EST	18-DEC-1999					
KEYWORDS	wf2d0d08.xl Soares Dieckgraefe.colon.NHUC Homo sapiens cDNA clone IMAGE:2351151.3 similar to TR:088826 088826 GDB-5 PROTEIN.;, mRNA sequence.									
ACCSSION	A1660957									
VERSION	A1660957.1	GI:4764540								
KEYWORDS	EST.									
SOURCE	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	1 (bases 1 to 734)									
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP).									
JOURNAL	Tumor Gene Index									
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert Length: 820 Std Error: 0.00 Seq primer: -400P from Glbco High quality sequence stop: 477. Location/Qualifiers 1..734 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2351151" /clone_1id="Soares_Dieckgraefe_colon_NHUC" /tissue_type="colononic mucosa from 5 ulcerative colitis patients" /lab_host="DH10B (phage-resistant)" /note="Organ: colon; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st transcribed cDNA was primed with Not I - oligo(dT) primer [5' TGTTACCACTGTGAAGTGAGGAGCGCGCGCATGCTTTTTTTTTTTTTTTPP 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieckelm.wustl.edu); colonnic mucosa represents a range of disease involvement from mild cryptitis to severe ulceration, fibrosis, and degeneration. Library constructed by Bento Soares and M. Fatima Bonaldo."									
FEATURES										
SOURCE										
BASE COUNT	200 a	143 c	159 g	228 t	4 others					
ORIGIN										
Query Match	100.0%; Score 154;	DB 10;	Length 734;							
Best Local Similarity	100.0%; Pred. No. 1e-36;									
Matches 154;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;						
OY	1	ACAGATCTTTTCATTGCGTATTCAGCGCTGTGGTAAGAAGTCGATCGAATAATCAGAAATATCC	60							
DB	261	ACAGATCTTTTCATTGCGTATTCAGCGCTGTGGTAAGAAGTCGATCGAATAATCAGAAATATCC	222							

```

1 61 AACATTGCACGAGTATCTTTGGTTATTCCTCCACAGACATCCGCCAGACACTGTACTCT 120
2 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
3 21 AACATTGCACGAGTATCTTTGGTTATTCCTCCACAGACATCCGCCAGACACTGTACTCT 162
4 121 GATGAACGCTGCTCCTTGCTGTCCTAATATTCATA 154
5 151 GATGAACGCTGCTCCTTGCTGTCCTAATATTCATA 128
6 RESULT 4 9
7 A1802756/c
8 LOCUS
9 DEFINITION
10 w18a04.x1 Soares_Dieckgreife.colon.NHUC Homo sapiens cDNA clone
11 IMAGE:2350926 3' similar to TR:088826 088826 GDB-5 PROTEIN.; mRNA
12 sequence.
13 ACCESSION
14 A1802756
15 VERSION
16 A1802756.1 GI:5368228
17 SOURCE
18 EST.
19 human:
20 HOMO SAPIENS
21 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
22 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
23 REFERENCE
24 1 (bases 1 to 561)
25 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
26 AUTHORS
27 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
28 TITLE
29 Tumor Gene Index
30 Unpublished (1997)
31 JOURNAL
32 COMMENT
33 Contact: Robert Strausberg, Ph.D.
34 Email: cgapbs-remail.nih.gov
35 This clone is available royalty-free through LNL ; contact the
36 IMAGE Consortium (infoimage.lnl.gov) for further information.
37 Insert length: 840 Std Error: 0.00
38 Seq primer: -40UP from Gibco
39 High quality sequence stop: 398.
40 FEATURES
41 source
42 Location/Qualifiers
43 1..561
44 /organism="Homo sapiens"
45 /db_xref="taxon:9606"
46 /clone="IMAGE:2350926"
47 /clone_lib="Soares_Dieckgreife.colon.NHUC"
48 /tissue_type="colonic mucosa from 5 ulcerative colitis
49 patients"
50 /lab_host="DH10B (phage-resistant)"
51 /note="Organ: colon; Vector: pTZ19-Pac (Pharmacia) with a
52 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
53 strand cDNA was primed with a Not I - oligo(dT) primer [5
54 TGTTACAATCTGAAGTGGGAGCGCGCTGATCTTTTTTTTTTTT 3'],
55 double-stranded cDNA was ligated to Eco RI adaptors
56 (Pharmacia), digested with Not I and cloned into the Not I
57 and Eco RI sites of the modified pTZ19 vector. Library
58 went through one round of normalization. Tissue samples
59 provided by Dr. Brian Dieckgreife (Washington University,
60 dieck@lm.wustl.edu); colonic mucosa represents a range of
61 disease involvement from mild cryptitis to severe
62 ulceration, fibrosis, and degeneration. Library
63 constructed by Bento Soares and M. Fatima Bonaldo. "
64 BASE COUNT
65 166 a 98 c 118 g 178 t 1 others
66 ORIGIN
67 Query Match 99.0%; Score 152.4; DB 10; Length 561;
68 Best Local Similarity 99.4%; Pred. No. 3e-36;
69 Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0.
70 1 ACAGATCTTTTCATTGCTATTCAGGCGCTGTTAAAGTGATCTGAATAACAGAATATCC 60
71 15 ACAGATCTTTTCATTGCTATTCAGGCGCTGTTAAAGTGATCTGAATAACAGAATATCC 238
72 61 AACATTGCACGAGTATCTTTGGTTATTCCTCCACAGACATCCGCCAGACACTGTACTCT 120
73 237 AACATTGCACGAGTATCTTTGGTTATTCCTCCACAGACATCCGCCAGACACTGTACTCT 178

```


OY	121	GATGAACGTCGTCTGCCTTGGCCCAAAATTTCATA	154
Db	177	GATGAACGTCGTCTGCCTTGGCCCAAAATTTCATA	144
RESULT	10		
		AM009764/c	EST
		LOCUS	08-MAR-2000
		DEFINITION	w887b05.xl NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:2504913 3'
		ACCESSION	Similar to TR:088826 O88826 GDB-5 PROTEIN.; mRNA sequence.
		VERSION	AM009764
		KEYWORDS	AM009764.1 GI:5858542
		SOURCE	EST.
		ORGANISM	human.
			Homo sapiens
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
			Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
		REFERENCE	1 (bases 1 to 450)
		AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
		TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
		JOURNAL	Unpublished (1997)
		COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-femail.nih.gov Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
			cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arraying: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/dbrr/image/image.html Insert Length: 587 Std Error: 0.00 Seq primer: -40UP from Gibco.
FEATURES			
Source			
		Location/Qualifiers	
		1..450	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone_image="IMAGE:2504913"	
		/clone_1lb="NCI_CGAP_Co3"	
		/sex="pooled"	
		/tissue_type="colon"	
		/lab_host="DH10B"	
		/note="vector: pRTD-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTD vector. Library went through one round of normalization."	
BASE COUNT	148 a	80 c	144 t
ORIGIN			
Query Match		98.1%; Score 151; DB 10; Length 450;	
Best Local Similarity	100.0%; Pred. No. 7.9e-36;		
Matches 151; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
OY	4	GATCTTTTCATTGCTAATCAGGCTGTTGTAAGTGATCTGAATCAGAAATATCCAAC	63
Db	450	GATCTTTTCATTGCTAATCAGGCTGTTGTAAGTGATCTGAATCAGAAATATCCAAC	391
OY	64	ATTGCACGAGTATCTTTGTTTATTCCTCCACAGACGCCGACGACTGCTCAT	123
Db	390	ATTGCACGAGTATCTTTGTTTATTCCTCCACAGACGCCGACGACTGCTCAT	331
OY	124	GAACGCTGCTCTCTGCTCTTATATTCATA	154
Db	330	GAACGCTGCTCTCTGCTCTTATATTCATA	300
RESULT	11		

	BGI95179	466 bp	mRNA	EST	21-APR-2001
LOCUS	BGI95179	Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.			
DEFINITION	BGI95179				
VERSION	BGI95179.1	GI:13718686			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
REFERENCE					
AUTHORS	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Dahl,T., Thornton,M., Ramchandran,K., Whittington,J.E, Lerner,L., Krashoc,D., Melligott,K., Clark,S., Mays,R., Smith,E. Veloso,N., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,D. and Ducat,M.				
TITLE	Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression				
JOURNAL	Nat. Biotechnol. 19 (5), 440 (2001) In press				
COMMENT:	Contact: Scott J. Cain Athersys, Inc. 3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900 Fax: 216 361 9596 Email: scaine@atersys.com High quality sequence stop: 466. Location/Qualifiers 1..466 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lbp="Athersys' RAGE Library" /cell_line="HT1080" /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression' Nature Biotechnology, in press. Note that even though 'the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."				
BASE COUNT	155 a 92 c 74 g 144 t				1 others
ORIGIN					
Query Match	Best Local Similarity	97.9%; Score 150.8; DB 11; Length 466; 98.7%; Pred. No. 9.2e-36;			
Matches 152; Conservative	0; Mismatches 2; Indels 0; Gaps				
QY	1 ACAGATCTTTCATTGCGTAACGGCTGTATTAAGTGCAATCGGAATAATCATCC 60				
DB	1B ACAGATCTTTCATTGCGTAATCTGCACAATGATTAAGGTCGATCGAAATATGCC 224				
OY	61 AACATTTGACGAGATATCTTTGTATATTCCTCCACAGAATCCGCCACAGACCTAGTCT 120				
DB	225 AACATTGCAAGATATCTTTGTATATTCCTCCACAGAATCCGCCACAGACCTAGTCT 284				
QY	121 GATGAACGCTGCTGCTGTTCCTTAATATTCATA 154				
DB	235 GATGAACGCTGCTGCTGTTCCTTAATATTCATA 318				
RESULT 12					
D25727	255 bp	mRNA	EST	30-NOV-1995	
LOCUS	HUMSGS04094 Human colon mucosa Homo sapiens cDNA clone cm2037 3'				
DEFINITION	mRNA sequence.				
ACCSSION	D25727				
VERSION	D25727.1	GI::500422			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
REFERENCE					
AUTHORS	Okubo,K., Yoshii,J., Yokouchi,H., Kameyama,M. and Matsubara,K.				

TITLE Global analysis of gene expression in colon mucosa: a large scale random cDNA sequencing analysis
 JOURNAL Unpublished (1994)
 COMMENT Contact: Okubo, K., Itoh, K., Yoshii, J., Yokouchi, H. and Matsubara, K. Institute for Molecular and Cellular Biology Osaka University
 3-1 Yamada-oka, Suita, Osaka 565, Japan.

FEATURES
 source
 1. .255
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="cm2037"
 /clone_lib="Human colon mucosa"
 /note="Adult male, tissue type = colon mucosa"
 76 a 52 c 39 g 80 t 8 others

BASE COUNT
 ORIGIN

Query Match 72.7%; Score 112; DB 11; Length 255;
 Best Local Similarity 97.4%; Pred. No. 5,8e-24;
 Matches 112; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 40 GATGTGAATGCAATATATCAACATATGACGAGTATCTTGTATTCCTCCACGACT 99
 |||||||
 DB 1 GATCTGAATGCAATATATCAACATATGACGAGTATCTTGTATTCCTCCACGACT 60
 |||||||

OY 100 CCGCCAGACACCTAGTCTGATGAACGCTGCTCTTGTCTTAATATTCATA 154
 |||||||
 DB 61 CCGCCAGACACCTAGTCTGATGAACGCTGCTCTTGTCTTAATATTCATA 115
 |||||||

RESULT 13
 A1687981 392 bp mRNA EST 27-MAY-1999
 LOCUS wa77f06.x1 Soares NFL_T.GBC.S1 Homo sapiens cDNA clone
 DEFINITION IMAGE:2302211 3' similar to TR-088826 088826 GGB-5 PROTEIN. ;, mRNA
 sequence.

ACCESSION
 VERSION A1687981 GI:4899275
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 1 (bases 1 to 392)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: 40UP from C1bco.

FEATURES
 source
 1. .392
 Location/Qualifiers
 1. .392
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2302211"
 /clone_lib="Soares.NFL_T.GBC.S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBH19W, testis NHT, and B-cell
 NCI-CGAP-GCB1) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo.

BASE COUNT 125 a 69 c 67 g 131 t
 ORIGIN

Query Match 59.0%; Score 90.8; DB 10; Length 392;
 Best Local Similarity 97.9%; Pred. No. 1,8e-17;
 Matches 92; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 131 AACATGACAGATATCTTGTATTCCTCCACAGCTCCGACAGACCTAGTCTCT 120
 |||||||
 DB 352 AACATGACAGATATCTTGTATTCCTCCACAGCTCCGACAGACCTAGTCTCT 333
 |||||||

OY 121 GATGAACGCTGCTCTTGTCTTAATATTCATA 154
 |||||||
 DB 332 GATGAACGCTGCTCTTGTCTTAATATTCATA 299
 |||||||

RESULT 14
 H1292301 405 bp mRNA EST 19-JUL-2001
 LOCUS UI-R-DNO-cliv-o-22-0-UI.s1 UI-R-DNO Rattus norvegicus cDNA clone
 DEFINITION UI-R-DNO-cliv-o-22-0-UI 3', mRNA sequence.

ACCESSION
 VERSION B1292301 GI:14952661
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 405)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the
 non-normalized distal colon library cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLA=yes.

FEATURES
 source
 1. .405
 Location/Qualifiers
 1. .405
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-DNO-cliv-o-22-0-UI"
 /clone_lib="UI-R-DNO"
 /dev_stage="ADULT"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DNO
 library is a non-normalized Rat distal colon library
 constructed in pT73 PAC vector according to the procedure
 described by Bonaldo, Lennon & Soares (Genome Research
 Genome 6: 791-806, 1996). The oligonucleotide used to
 prime first strand synthesis contained the sequence tag
 CCAGG between the Not I cloning site and dT18 stretch.
 TAG LIB=UI-R-DNO
 TAG_TISSUE=distal colon
 TAG_SEQ=GAAGTCTCC.

BASE COUNT	103 a	91 c	89 g	122 t
ORIGIN				

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:29 ; Search time 16681 Seconds

(without alignments)
140.434 Million cell updates/sec

Title: US-09-049-696-9

Perfect score: 1 GAACATGCGCTCATTTGATG.....TCTCTGGGATCCAGTCGAC 218

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 segs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_est1:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_tod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	218	100.0	625	10	AM361532 QV2-CT026
2	216.4	99.3	561	10	AM361523 QV2-CT026
3	216.4	99.3	621	10	AM361520 QV2-CT026
4	216.4	99.3	652	10	AM753449 QV2-CT026
5	216.4	99.3	654	10	AM361522 QV2-CT026
6	216.4	99.3	657	10	AM361521 QV2-CT026
7	156.2	71.7	455	10	AA871197 VQ3C11.r
8	156.2	71.7	661	11	BF581041 602100173
9	156.2	71.7	2915	12	AK007466 Mus muscu
10	156.2	71.7	2933	12	AK008659 Mus muscu
11	149.6	68.6	906	11	BF581342 602100726
12	135.4	62.1	550	11	BF353533 RC3-HT086

C	13	133.8	61.4	534	11	BF355513 RC3-HT086
C	14	130.6	59.9	935	11	BF578919 602095492
C	15	123.4	56.6	308	10	AV077718 AV077718
C	16	115	52.8	633	10	AM753451 QV2-CT026
C	17	103.4	47.4	653	10	AM361534 QV2-CT026
C	18	88.8	40.7	747	11	BF580957 602100659
C	19	82.6	37.9	510	10	AM762590 ur6408.y
C	20	82.6	37.9	584	11	BF116637 uy9907.y
C	21	82.6	37.9	866	11	BI155803 602904376
C	22	82.6	37.9	1033	11	BI243305 602355738
C	23	80	36.7	742	13	BH030431 RPT1-24-3
C	24	78.6	36.1	821	11	BG286204 602383026
C	25	78.4	36.0	462	10	AI851265 UI-M-BH0
C	26	76.8	35.2	235	10	BB299286 BB299286
C	27	76.2	35.0	704	11	BG965023 602831668
C	28	72.8	33.4	435	10	AA443258 aa14d09.r
C	29	72.8	33.4	458	11	BG675746 602621972
C	30	72.8	33.4	940	11	BG678187 602625970
C	31	71.8	32.9	356	10	AM361528 QV2-CT026
C	32	71.8	32.9	430	10	AI493356 t970f11.x
C	33	71	32.6	285	10	BB226010 BB226010
C	34	70	32.1	300	10	BB560274 BB560274
C	35	69.4	31.8	914	11	BG172513 602333711
C	36	66.8	30.6	257	10	BB560270 BB560270
C	37	66.8	30.6	314	10	BB225759 BB225759
C	38	66.6	30.6	279	10	BB475477 BB475477
C	39	65.6	30.1	299	10	BB124752 BB124752
C	40	65.6	30.1	299	10	BB500238 BB500238
C	41	64.6	29.6	280	10	BB043672 BB043672
C	42	61.8	28.3	331	10	BB499841 BB499841
C	43	59.4	27.2	301	13	AZ779200 ZM0015713
C	44	58.6	26.9	142	11	BF333499 RC4-CN000
C	45	56.4	25.9	437	10	BE382180 601271754

ALIGNMENTS

RESULT 1
LOCUS AM361532 625 bp mRNA
DEFINITION QV2-CT0261-261099-011-e03 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361532
VERSION AM361532.1 GI:6866286
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 625)
HCGP <http://www.judwig.org.br/ORESTES>.

JOURNAL
TITLE The FAPESP/LICR Human Cancer Genome Project
UNPUBLISHED (1999)
CONTACT: Simpson A.J.G.
LABORATORY OF Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

TELE: +55-11-2704922
FAX: +55-11-2707001
Email: asimpson@judwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.judwig.org.br/scripts/gethtml2.pl?rl=QV2&rl=QV2-CT0261-261099-011-e03&rl=1999-10-26&rl=1>)
Seq primer: puc 18 forward
High quality sequence start: 93
Location/Qualifiers
1..625

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"


```

/dev_stage="Adult"
/Note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
RNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      172 a      154 c      150 g      149 t
ORIGIN

```

```

Query Match      100.0%; Score 218; DB 10; Length 625;
Best Local Similarity 100.0%; Pred. No. 1.5e-56;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAATGGCCCTCATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGCTGTCTCTCA 60
DB 242 GAACAATGGCCCTCATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGCTGTCTCTCA 301
QY 61 GCGCTCCATCCAGCTTGAGAGATTAAGGATTAACCTCCAGAACGCCAGTGAATG 120
DB 302 GCGCTCCATCCAGCTTGAGAGATTAAGGATTAACCTCCAGAACGCCAGTGAATG 361
QY 121 CACAGTATCGTGAGACAGCAGCCGTGGAAAGACACTTTGTTTATCAGCTGACAC 180
DB 362 CACAGTATCGTGAGACAGCAGCCGTGGAAAGACACTTTGTTTATCAGCTGACAC 421
QY 181 GCAGCTCCGCCAAATCCTCTCTGGGATCCAGTGAC 218
DB 422 GCAGCTCCGCCAAATCCTCTCTGGGATCCAGTGAC 459

```

```

RESULT 2
AM361523      561 bp      mRNA      EST      04-FEB-2000
LOCUS      OV2-CT0261-261099-011-g05 CT0261 Homo sapiens cDNA, mRNA sequence.
DEFINITION      AM361523
ACCESSION      AM361523.1 GI:6866277
VERSION      EST.
KEYWORDS      human.
SOURCE      Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 561)
AUTHORS      HCGP http://www.ludwig.org.br/ORESTES.
TITLE      The FAPESP/LICR Human Cancer Genome Project
JOURNAL      Unpublished (1999)
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=QV2&l2=QV2-CT0261-
261099-011-g05&l3=1999-10-26&l4=1)
Seq primer: puc 18 forward
High quality sequence start: 43
High quality sequence stop: 560.
Location/Qualifiers
1. 561
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/Note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
RNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      166 a      151 c      150 g      154 t
ORIGIN

```

FEATURES

```

source
1. 561
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/Note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
RNA and cDNA amplification were performed under low
stringency conditions."

```

```

into the pUC 18 vector. Reverse transcription of tissue
RNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      154 a      140 c      140 g      127 t
ORIGIN

```

```

Query Match      99.3%; Score 216.4; DB 10; Length 561;
Best Local Similarity 99.5%; Pred. No. 4.4e-56;
Matches 217; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAACAATGGCCCTCATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGCTGTCTCTCA 60
DB 196 GAACAATGGCCCTCATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGCTGTCTCTCA 255
QY 61 GCGCTCCATCCAGCTTGAGAGATTAAGGATTAACCTCCAGAACGCCAGTGAATG 120
DB 256 GCGCTCCATCCAGCTTGAGAGATTAAGGATTAACCTCCAGAACGCCAGTGAATG 315
QY 121 CACAGTATCGTGAGACAGCAGCCGTGGAAAGACACTTTGTTTATCAGCTGACAC 180
DB 316 CACAGTATCGTGAGACAGCAGCCGTGGAAAGACACTTTGTTTATCAGCTGACAC 375
QY 181 GCAGCTCCGCCAAATCCTCTCTGGGATCCAGTGAC 218
DB 376 GCAGCTCCGCCAAATCCTCTCTGGGATCCAGTGAC 413

```

```

RESULT 3
AM361520      621 bp      mRNA      EST      04-FEB-2000
LOCUS      OV2-CT0261-261099-011-e07 CT0261 Homo sapiens cDNA, mRNA sequence.
DEFINITION      AM361520
ACCESSION      AM361520.1 GI:6866274
VERSION      EST.
KEYWORDS      human.
SOURCE      Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 621)
AUTHORS      HCGP http://www.ludwig.org.br/ORESTES.
TITLE      The FAPESP/LICR Human Cancer Genome Project
JOURNAL      Unpublished (1999)
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=QV2&l2=QV2-CT0261-
261099-011-e07&l3=1999-10-26&l4=1)
Seq primer: puc 18 forward
High quality sequence start: 88
High quality sequence stop: 613.
Location/Qualifiers
1. 621
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/Note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
RNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      166 a      151 c      150 g      154 t
ORIGIN

```

FEATURES

```

source
1. 621
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/Note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
RNA and cDNA amplification were performed under low
stringency conditions."

```


Query Match 99.3%; Score 216.4; DB 10; Length 621;
 Best Local Similarity 99.5%; Pred. No. 4.5e-56;
 Matches 217; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAACAATGGCCCTCATGATGATGCTTTGGGGCCCTTCATCAGAAATGACCTGCTCTCA 60
 DB 240 GAACAATGGCCCTCATGATGATGCTTTGGGGCCCTTCATCAGAAATGACCTGCTCTCA 299
 QY 61 GGGCTCATTCAGCTTGAGAGTAAGGATTAACCTCCAGAAACGACGATGAATGG 120
 DB 300 GGGCTCATTCAGCTTGAGAGTAAGGATTAACCTCCAGAAACGACGATGAATGG 359
 QY 121 CACAGTATGCTGGACAGCAGCCGTGGAAAGACACTTTGTTCTTATCAGCTGGACAC 180
 DB 360 CACAGTATGCTGGACAGCAGCCGTGGAAAGACACTTTGTTCTTATCAGCTGGACAA 419
 QY 181 GCAGCCTCCCAATCCTTCTCTGGAGATCCAGTGAGC 218
 DB 420 GCAGCCTCCCAATCCTTCTCTGGAGATCCAGTGAGC 457

RESULT 4
 AW753449 652 bp mRNA EST 28-APR-2000
 LOCUS AW753449
 DEFINITION QV2-CT0261-261099-011-e08 CT0261 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW753449
 VERSION AW753449.1 GI:7668381
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 652)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.G.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL CONTACT: Simpson A.J.G.
 MEDLINE Laboratory of Cancer Genetics
 COMMENT Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-CT0261-
 261099-011-e08&t3=1999-10-26&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 14.

FEATURES
 source 1..652
 Location/Qualifiers
 1..652
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0261"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
 Sma1; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 179 a 153 c 161 g 159 t

Query Match 99.3%; Score 216.4; DB 10; Length 652;
 Best Local Similarity 99.5%; Pred. No. 4.0e-56;
 Matches 217; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 61 GAACAATGGCCCTCATGATGATGCTTTGGGGCCCTTCATCAGAAATGACCTGCTCTCA 60
 DB 240 GAACAATGGCCCTCATGATGATGCTTTGGGGCCCTTCATCAGAAATGACCTGCTCTCA 359
 QY 61 GGGCTCATTCAGCTTGAGAGTAAGGATTAACCTCCAGAAACGACGATGAATGG 120
 DB 300 GGGCTCATTCAGCTTGAGAGTAAGGATTAACCTCCAGAAACGACGATGAATGG 419
 QY 121 CACAGTATGCTGGACAGCAGCCGTGGAAAGACACTTTGTTCTTATCAGCTGGACAC 180
 DB 360 CACAGTATGCTGGACAGCAGCCGTGGAAAGACACTTTGTTCTTATCAGCTGGACAA 479
 QY 181 GCAGCCTCCCAATCCTTCTCTGGAGATCCAGTGAGC 218
 DB 420 GCAGCCTCCCAATCCTTCTCTGGAGATCCAGTGAGC 517

RESULT 5
 AW361522 654 bp mRNA EST 04-FEB-2000
 LOCUS AW361522
 DEFINITION QV2-CT0261-261099-011-f10 CT0261 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW361522
 VERSION AW361522.1 GI:6866276
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 654)
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 HCGP http://www.ludwig.org.br/ORESTES.
 The FAPESP/LICR Human Cancer Genome Project
 Unpublished (1999)
 JOURNAL CONTACT: Simpson A.J.G.
 COMMENT Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-CT0261-
 261099-011-f10&t3=1999-10-26&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 132
 High quality sequence stop: 643.

FEATURES
 source 1..654
 Location/Qualifiers
 1..654
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0261"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
 Sma1; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 173 a 156 c 162 g 163 t

Query Match 99.3%; Score 216.4; DB 10; Length 654;
 Best Local Similarity 99.5%; Pred. No. 4.0e-56;

Matches 217: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAACAATGGCCATCATGCTTTGGGGCCCTTCATCAGAAATGAGAGCTGTCTCA 60
|||||
Db 292 GAACAATGGCCATCATGCTTTGGGGCCCTTCATCAGAAATGAGAGCTGTCTCA 351
|||||

QY 61 GCGCTCCATCCAGCTTGAAGATTAACCTCCAGAACAGCCAGTGAATG 120
|||||
Db 352 GCGCTCCATCCAGCTTGAAGATTAACCTCCAGAACAGCCAGTGAATG 411
|||||

QY 121 CACAGTATCGTGACAGACAGCCGGAAGACACTTGTCTTATCAGCTGACAC 180
|||||
Db 412 CACAGTATCGTGACAGACAGCCGGAAGACACTTGTCTTATCAGCTGACAT 471
|||||

QY 181 GCAGCCTCCCAATTCCTCTCTGGATCCAGTGAC 218
|||||
Db 472 GCAGCCTCCCAATTCCTCTCTGGATCCAGTGAC 509
|||||

RESULT 6
AM361521 657 bp mRNA EST 04-FEB-2000
LOCUS OY2-CT0261-261099-011-F03 CT0261 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION AM361521
VERSION AM361521.1 GI:6866275
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 657)
AUTHORS HCGP <http://www.ludwig.org.br/ORESTES>.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?rl=OY2-CT0261-261099-011-f03&f3=1999-10-26&f4=1>)
Seq primer: puc 18 forward
High quality sequence start: 123
High quality sequence stop: 656.
Location/Qualifiers
1. 657
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 175 a 160 c 164 g 158 t
ORIGIN

Query Match 99.3%; Score 216.4; DB 10; Length 657;
Best Local Similarity 99.5%; Pred. No. 4.6e-56;
Matches 217; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAACAATGGCCATCATGCTTTGGGGCCCTTCATCAGAAATGAGAGCTGTCTCA 60
|||||
Db 288 GAACAATGGCCATCATGCTTTGGGGCCCTTCATCAGAAATGAGAGCTGTCTCA 347
|||||

QY 31 GCGCTCATCCAGCTTGAAGATTAACCTCCAGAACAGCCAGTGAATG 120
|||||
Db 318 GCGCTCATCCAGCTTGAAGATTAACCTCCAGAACAGCCAGTGAATG 407
|||||

QY 121 CACAGTATCGTGACAGACAGCCGGAAGACACTTGTCTTATCAGCTGACAC 180
|||||
Db 408 CACAGTATCGTGACAGACAGCCGGAAGACACTTGTCTTATCAGCTGACAT 467
|||||

QY 181 GCAGCCTCCCAATTCCTCTCTGGATCCAGTGAC 218
|||||
Db 468 GCAGCCTCCCAATTCCTCTCTGGATCCAGTGAC 505
|||||

RESULT 7
AA871197 455 bp mRNA EST 16-MAR-1998
LOCUS vq32c11.1 Barstead bowel MPLB9 Mus musculus cDNA clone
DEFINITION IMAGE:1095956 5' similar to SW:ECIC_BOVIN P54281 EPITHELIAL
CHLORIDE CHANNEL PROTEIN ;, mRNA sequence.
ACCESSION AA871197
VERSION AA871197.1 GI:2966642
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 455)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucab,T., Lacy,M., Le,M., Martin,T., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The Washu-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washu-HMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:602188
Seq primer: -28m13 rev2 EF from Amersham
High quality sequence stop: 414.
Location/Qualifiers
1. 455
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1095956"
/clone_lib="Barstead bowel MPLB9"
/tissue_type="bowel"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pUT30-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTACGATCTGAAGAGGAGCGCGCCCTTTTATTTTATTTTATTTTATTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AATTCGATCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pUT30 vector.
Source irradiated bowel harvested 72 hours after
irradiation (1400 Gys). Library constructed by Bob
Barstead.
BASE COUNT 122 a 123 c 105 g 105 t
ORIGIN

Query Match 71.7%; Score 156.2; DB 10; Length 455;
Best Local Similarity 82.5%; Pred. No. 1.3e-37;

Matches 179; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1 GAACAATGCGCTCATATGCTTTGGGCGCTTTCATCAGGAATGAGCTGTCTCA 60
D 4 GAACAATGCTGTGTGATGCTTTCAGCACTCTCTCAGGAATGGCGCATGCTCA 63
QY 61 GCGCTCCATCCAGCTTGAAGTAAAGGATTAACCTCCAGAACAGCCAGTGAATGG 120
D 64 GCATCCATCCAGCTGAGAGAGAGAGAGAGATTAATCTCAGATTAACATGATGAATGG 123
QY 121 CACAGTGTCTGGACAGCAGCGGGAAGAGACCTTCTTATCACTGAGACAC 180
D 124 CTCAGTGTCTGGACAGCAGCGGGAAGAGACCTTCTTATCACTGAGACAC 183
QY 181 GCAGCTCCCGCAATCCCTCTCTGGATCCAGTGA 217
D 184 GCATCTCTCTCAATATTTATCTGGATCCAGCGGA 220

RESULT 8
BF581041 661 bp mRNA EST 12-DEC-2000
LOCUS 602100173P1 NCI_CGAP_C024 Mus musculus cDNA clone IMAGE:4219698 5',
DEFINITION mRNA sequence.
ACCESSION BF581041
VERSION BF581041 GI:11654753
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 661)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gspbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9802 row: b column: 19
High quality sequence stop: 645.
location/Qualifiers
1. 661
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4219698"
/clone_lib="NCI_CGAP_C024"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: colon; Vector: pCMV-Sport6; Site: 1; Not:
Site: 2; Salt: Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 180 a 160 c 172 g 149 t
ORIGIN

Query Match 71.7%; Score 156.2; DB 11; Length 661;
Best Local Similarity 82.5%; Pred. No. 1.4e-37;
Matches 179; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1 GAACAATGCGCTCATATGCTTTGGGCGCTTTCATCAGGAATGAGCTGTCTCA 60
D 361 GAACAATGCTGTGTGATGCTTTCAGCACTCTCTCAGGAATGGCGCATGCTCA 420
QY 61 GCGCTCCATCCAGCTTGAAGTAAAGGATTAACCTCCAGAACAGCCAGTGAATGG 120
D 421 GCATCTCTCTCAATATTTATCTGGATCCAGCGGA 480

QY 121 CACAGTGTCTGGACAGCAGCGGGAAGAGACCTTCTTATCACTGAGACAC 180
D 481 CTCAGTGTCTGGACAGCAGCGGGAAGAGACCTTCTTATCACTGAGACAC 540
QY 181 GCAGCTCCCGCAATCCCTCTCTGGATCCAGTGA 217
D 541 GCATCTCTCTCAATATTTATCTGGATCCAGCGGA 577

RESULT 9
AK007466 2915 bp mRNA HTC 05-JUL-2001
LOCUS Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
enriched library, clone:1810012P03, full insert sequence.
ACCESSION AK007466
VERSION AK007466 GI:12841032
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to
clone:1810012P03.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2915)
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLES High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2 (bases 1 to 2915)
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome research. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED 11076861
3 (bases 1 to 2915)
Shibata, K., Itoh, M., Aizawa, K., Nagoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384 format
sequencing pipeline with 384 multicapillary sequencer
Genome research. 10 (11), 1757-1771 (2000)
JOURNAL 20530913
MEDLINE 11076861
PUBMED 11076861
4 (bases 1 to 2915)
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 2915)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,
Hara, K., Hara, A., Hayatsu, N., Hiramoto, K., Hirata, T., Horii, F.,
Imolandi, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,
Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,
Miyazaki, A., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Okazaki, Y.,
Okido, T., Owa, C., Saito, R., Saito, R., Sakai, C., Sakai, K., Sano, H.,
Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

CDS

```

/tissue_type="stomach"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
29..2770
/note="putative"
/codon_start=1
/protein_id="BAB25815.1"
/db_xref="GI:12842988"
/translation="MESLSPVFLILHLEGLSESLIQLNNNGEIVIAIDHDV
EDALIHKIDKQASPYLFGATGKREFKVALIPESYAKPEYPTKETRNA
DIVSTSPDGDEPDETHGACGEKRIHLTPPLAGKLTQYGPDRVHEMAH
FRMGVFNNDDEKYLKSKQAVRCSATLGAQVRRGCGSCITNKCVIDRTG
LYKDCVFPDPHNEKASIMPNONINSVECTGKNHNEAPNODNOCNLRSTEW
IOESDEKQTPMTAPAPPTSLQIQRIQVCLVLDKSGMLNDRMLRMQASLFL
LQVDEQSGVMWTFDSAAVQSEIKOLNSGADRLIKHLPTVSAGTSICSLRT
AFVTKRKYPTDSEIVLITDEGNTISSCPDLVOSGAIHTVALGPAKAELEQL
KMTGLQYSSDOVNGIIVAFALSSGMAIACHSTOLESGYNLONNOMNKSLSI
VDSVSGKTLTLITWTPPTFFIMDPGSEVNGTLDITTVATLQVFGTKVGEFW
YSIDASSQTLITVTSRAASATLPITVTPVKNNTGKPSPTVYASIRQASPLR
ASVVALIESVNGKTVTLLELNGAGADATKNDGVSRFPAPDANGRSVKTALGCV
TSDRAPAPKPNRPYIDGWIEDGEVRNPPPEYVODKOLCFSSSGSFVATN
VPAAPIDPLEPPCOITDLKASIOGONVNLWTPAGDEYDHGRASNTIRMSSTIVD
LRHNTSLQVNTGLIPKEASSEEIPEFEIGNGTDFGNGTDFIATQVADKSNLSEI
SNIAVSFPIAPQEPPIEDSPPCDISINSTIGIHVLIIMKMLGMQVTLGLH"
2912..2917
/note="putative"
2933
polya_site
/note="putative"
2933
BASE COUNT      845 a      722 c      697 g      669 t
ORIGIN

```

```

Query Match      71.7%; Score 156.2; DB 12; Length 2933;
Best Local Similarity 82.5%; Pred. No. 2e-37;
Matches 179; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```

```

QY 1 GAACAATGCCCTCATGATGCTTTTGGGCCCTTCATCAGCAATGAGCTGTCTCTCA 60
    |||||  ||  |||||  ||  ||  |||||  ||  ||  |||||  ||  ||  |||||
DB 1423 GAACAATGGCTGTGTGATGCTTTTCACACACTCTCTCAGCAATGAGCGATCGCTCA 1482
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 61 GCGCTCATCCAGCTTGAGAGTAAGGATTAACCTCCAGCAAGCAGCATGATGATG 120
    ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1483 GCACTCATCCAGCTGAGAGGAGGAGGAGTATATCTCCAGCAATGAGATGATGATG 1542
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 CACACTGATCTGAGACAGCAGCCGCGGGAAGAGACACTTGTCTTATCAGCTGAGCAAC 180
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1543 CTCACATGANTCGGAGACAGCTGCGTGGGCAAGACACTGTCTTATCAGCTGAGCAAC 1602
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 181 GCAGCCTCCCAATCTCTCTCTGAGATCCAGTGA 217
    ||  |||||  ||  ||  |||||  |||||  |||||  |||||  |||||
DB 1603 GCATCTCTCTACAAATTTATCTGGATCCAGCGGA 1639
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

RESULT 11
LOCUS      BF581342      906 bp      mRNA      EST      12-DEC-2000
DEFINITION 602100726r1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4223935 5',
            mRNA sequence.
ACCESSION  BF581342
VERSION     BF581342.1 GI:11655054
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 906)
            NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
AUTHORS    Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-r@mail.nih.gov
COMMENT    Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNU)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNU at:
<http://image.llnl.gov>
 Plate: LLAM9813 row: c column: 08
 High quality sequence stop: 697.
 Location/Qualifiers

```

source          1..906
                /organism="Mus musculus"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone_image="4223935"
                /clone_lib="NCI_CGAP_Co24"
                /lab_host="DH10B (71 phage-resistant)"
                /note="Organ: colon; Vector: pCMV-Sport6; Site_1: NciI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.6 Kb. Constructed by Life
                Technologies. Note: this is a NCI_CGAP library."
BASE COUNT      251 a      237 c      246 g      172 t
ORIGIN

```

```

Query Match      68.6%; Score 149.6; DB 11; Length 906;
Best Local Similarity 81.6%; Pred. No. 1.6e-35;
Matches 173; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```

```

QY 6 ATGGCTCATGATGCTTTGGGGCCCTTCATCAGCAATGAGCTGTCTCAGCGCT 65
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 192 AAGCTCTGTGTGATGCTTTCGACGACCTCTCAGCAATGAGCGCTCAGACT 91
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 66 CCATCCAGCTTGAGAGTAAGGATTAACCTCCAGCAAGCAGCATGATGATGAGCAG 125
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 92 CCATCCAGCTTGAGAGTAAGGATTAACCTCCAGCAATGATGATGATGATGATG 151
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 126 TGATCGTGGACACACCGCTGGGAAAGAGCACTTGTCTTATCAGCTGAGCAACGAGC 185
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 192 TGATCGTGGACACCGCTGGGCAAGAGCACCTTGTCTTATCAGCTGAGCAACGAGC 211
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 186 CTCGCCAATCTCTCTCTGAGATCCAGTGA 217
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 212 CTCCTACAAATTTATCTGGATCCAGCGGA 243
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

RESULT 12
LOCUS      BF355533      550 bp      mRNA      EST      22-NOV-2000
DEFINITION RC3-H20865-030800-021-b01 HT0865 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF355533
VERSION     BF355533.1 GI:11314594
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 550)
            Dias Neto, E., da Silva, W. Jr., Zagdo, M.A., Bordin, S., Costa, P.F.,
            Nagai, M.A., de Almeida, R., Varjovski-Almeida, S., Briones, M.R.,
            Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
            Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
            M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil.
Tel: +55-11-2704922
Fax: +55-11-2707001

```


Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&ct=RC3-HR0865-030800-021-p01&t3=2000-08-03&t4=1>)

Seq primer: puc 18 forward
 High quality sequence start: 71
 High quality sequence stop: 548
 Location/Qualifiers

FEATURES

SOURCE

1..550
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="HR0865"
 /dev_stage="Adult"
 /note="Organ: head,neck; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 148 a 111 c 127 g 164 t
 ORIGIN

Query Match 62.1%; Score 135.4; DB 11; Length 550;
 Best local similarity 76.5%; Pred. No. 3.3e-31;

Matches 166; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 1 GAACAATGGCCCTCATGTATGCTTTGGGCGCCCTTCATCAGGAATGAGCTGCTCTCA 60
 DB 511 GAACAATGGCCCTCATGTATGCTTTGGGCGCCCTTCATCAGGAATGAGCTGCTCTCA 452
 QY 61 GCGCTCCATCCAGCTTGAGATTAAGGATTAACCTCCAGAACACCCAGTGGATTAATGG 120
 DB 451 GAAGTCCCTTACGCTGCGAAGATTAAGGATTAACCTGATTAATGCTGGATTAACGA 392
 QY 121 CACAGTATCGTGGACACACCGCTGGAAGGACACTTTGTTCTTATCAGCTGGACAAAC 180
 DB 391 CACGTCTATTAATGATTAATACAGTGGGAAGACAGCTTTCTTCAATCAGATGAGACG 332
 QY 181 GCAGCTCCCAAAATCCTTCTGTGGATCCAGTGA 217
 DB 331 TCTGCTCCCAAGTATTTCTCTGTGGATCCAGTGA 295

RESULT 13

BF355513/c

LOCUS 534 bp mRNA EST 22-NOV-2000
 DEFINITION RC3-HR0865-030800-021-c08 HR0865 Homo sapiens CDNA, mRNA sequence.
 ACCESSION BF355513
 VERSION BF355513.1 GI:11314600

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 534)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W.J.F., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&ct=RC3-HR0865-030800-021-p01&t3=2000-08-03&t4=1>)

Seq primer: puc 18 forward
 High quality sequence stop: 534.
 Location/Qualifiers

FEATURES

SOURCE

1..534
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="HR0865"
 /dev_stage="Adult"
 /note="Organ: head,neck; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 145 a 110 c 118 g 161 t
 ORIGIN

Query Match 61.4%; Score 133.8; DB 11; Length 534;
 Best local similarity 76.0%; Pred. No. 1e-30;

Matches 165; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1 GAACAATGGCCCTCATGTATGCTTTGGGCGCCCTTCATCAGGAATGAGCTGCTCTCA 60
 DB 499 GAACAATGGCCCTCATGTATGCTTTGGGCGCCCTTCATCAGGAATGAGCTGCTCTCA 440
 QY 61 GCGCTCCATCCAGCTTGAGATTAAGGATTAACCTCCAGAACACCCAGTGGATTAATGG 120
 DB 499 GAAGTCCCTTACGCTGCGAAGATTAAGGATTAACCTGATTAATGCTGGATTAACGA 380
 QY 121 CACAGTATCGTGGACACACCGCTGGAAGGACACTTTGTTCTTATCAGCTGGACAAAC 180
 DB 379 CACGTCTATTAATGATTAATACAGTGGGAAGACAGCTTTCTTCAATCAGATGAGACG 320
 QY 181 GCAGCTCCCAAAATCCTTCTGTGGATCCAGTGA 217
 DB 319 TCTGCTCCCAAGTATTTCTCTGTGGATCCAGTGA 283

RESULT 14

BF578919

LOCUS 935 bp mRNA EST 12-DEC-2000
 DEFINITION 602095492P1 NCI_CGAP_CO24 Mus musculus CDNA clone IMAGE:4215598 5',
 ACCESSION BF578919
 VERSION BF578919.1 GI:11652631

KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 935)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA sequencing by: Incyte Genomics, Inc.
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: L1AM9791 row: 9 column: 23

(ojpsn) yuolq ednJ
Page Blank

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:53:49 ; Search time 521.98 Seconds
(without alignments)
94.586 Million cell updates/sec

Title: US-09-049-696-9
218
Sequence: 1 GACCAATGGCCTCATGTATG.....TCTCTGGATCCAGTGAC 218

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/pdata/1/lna/5A.COMB.seq:*
2: /cgn2_6/pdata/1/lna/5B.COMB.seq:*
3: /cgn2_6/pdata/1/lna/5A.COMB.seq:*
4: /cgn2_6/pdata/1/lna/5B.COMB.seq:*
5: /cgn2_6/pdata/1/lna/PCTUS.COMB.seq:*
6: /cgn2_6/pdata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29.8	13.7	5455	1	US-08-342-930-1
2	29.4	13.5	464	2	US-08-449-287-11
3	29.4	13.5	777	2	US-08-860-882A-28
4	28.6	13.1	603	4	US-09-411-329C-2
5	28.6	13.1	609	4	US-09-411-329C-6
6	28.6	13.1	816	4	US-09-411-329C-19
7	28.6	13.1	1373	4	US-09-411-329C-20
8	28.6	13.1	1386	4	US-09-411-329C-4
9	28.6	13.1	1392	4	US-09-411-329C-7
10	28.6	13.1	1620	4	US-09-411-329C-12
11	28.6	13.1	1620	4	US-09-411-329C-13
12	27.8	12.8	6463	2	US-08-962-284-3
13	27.6	12.7	464	2	US-08-449-287-9
14	27.6	12.7	632	4	US-09-328-111-53
15	26.8	12.3	371	2	US-08-602-716A-3
16	26.4	12.1	28804	3	US-08-592-874-1
17	26.4	12.1	28804	3	US-09-096-942-2
18	26.4	12.1	28804	3	US-09-096-942-2
19	26.2	12.0	2277	1	US-08-369-796-7
20	26.2	12.0	2277	2	US-08-852-091-7
21	26.2	12.0	2277	2	US-08-820-754-7
22	26.2	12.0	2277	3	US-08-956-652-7
23	26.2	12.0	2277	3	US-08-956-869-7
24	26.2	12.0	2277	3	US-08-948-547-7
25	26.2	12.0	2374	5	PCT-US95-17025-7
26	26.2	12.0	2374	5	US-08-466-589-5
27	26.2	12.0	2374	2	US-08-700-636-5

28	26.2	12.0	2374	3	US-08-467-574-5	Sequence 5, Appl1
29	26.2	12.0	2374	4	US-09-217-945-5	Sequence 5, Appl1
30	26	11.9	688	4	US-09-027-381-1	Sequence 1, Appl1
31	26	11.9	728	4	US-09-276-531-128	Sequence 128, App
32	25.4	11.7	139	5	PCT-US93-06251-158	Sequence 58, Appl
33	25.4	11.7	275	5	PCT-US93-06251-146	Sequence 46, Appl
34	25.4	11.7	596	5	PCT-US93-06251-152	Sequence 52, Appl
35	25.4	11.7	651	2	US-08-546-712-1	Sequence 1, Appl1
36	25.4	11.7	651	2	US-08-751-105-1	Sequence 1, Appl1
37	25.4	11.7	43676	3	US-09-356-952-12	Sequence 12, Appl
38	25.4	11.7	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
39	25.4	11.7	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
40	25.2	11.6	348	2	US-08-273-146-54	Sequence 54, Appl1
41	25.2	11.6	1410	4	US-09-255-368-1	Sequence 1, Appl1
42	25.2	11.6	1670	3	US-09-026-482B-1	Sequence 1, Appl1
43	25.2	11.6	2291	2	US-08-476-062A-53	Sequence 53, Appl
44	25.2	11.6	2310	2	US-08-476-062A-41	Sequence 41, Appl
45	25.2	11.6	2310	5	PCT-US96-01314-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-08-342-930-1
Sequence 1, Application US/08342930
Patent No. 5821084

GENERAL INFORMATION:

APPLICANT: OLIMSTED, ELIZABETH A.

APPLICANT: MAURO, LAURA J.

APPLICANT: DAVIS, ALAN R.

APPLICANT: DIXON, JACK E.

TITLE OF INVENTION: OSTEOBLAST-TESTICULAR PROTEIN TYROSINE

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/342,930

FILING DATE: 21-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KOSKI, ANTOINETTE F.

REGISTRATION NUMBER: 34,202

REFERENCE/DOCKET NUMBER: 20344-20975.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5455 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

FEATURE:

NAME/KEY: CDS

LOCATION: 205..5337

US-08-342-930-1

Query/Match 13.7%; Score 29.8; DB 1; Length 5455;
Best/Local Similarity 51.1%; Pred. No. 0.98;

Oy 36 CATCGGAAATGAGACTGCTGCTCCTCAGCCCTCATCTCAGCTTGAGAGCTAAAGGATTAACCC 95
 Db 191 CACCTGGAAAGGAGATCTCAGTGGCTGGGGCTTCATCGAAATPAGGCAATATGGATACACAA 25
 Oy 96 TCCAGAACAGCCACGATGATGAATGGCACAGATGATCGTGAGACAGCCCGTGGGAAAGGACA 15
 Db 251 CAGACTACTCTGCGATCTGTGAGAGGAAACATTCACACTTTCCAGAGACAAGACCAAGTCCA 31
 Oy 156 GTTGTGTTCTTATCACCTGACCAAGCAGCC 186
 Db 311 CACGTACTCTGAGATGAATACACGCGAGCC 341

```

RESULT 4
US-09-411-329C-2/c
: Sequence 2, Application US/09411329C
: Patent No. 6261820
: GENERAL INFORMATION:
: APPLICANT: Boone, Thomas
: APPLICANT: Li, HuiMin
: APPLICANT: Mann, Michael
: TITLE OF INVENTION: FIBRINOLYTICALLY ACTIVE POLYPEPTIDE
: FILE REFERENCE: A-596
: CURRENT APPLICATION NUMBER: US/09/411,329C
: CURRENT FILING DATE: 1999-10-01
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn version 3.0 .
: SEQ ID NO 2
: LENGTH: 603
: TYPE: DNA
: ORGANISM: Artificial
: FEATURE:
: OTHER INFORMATION: Native pro-NAT (analog of fibrolase)
: US-09-411-329C-2

```

Query Match	13.1%;	Score 28.6;	DB 4;	Length 603;
Best Local Similarity	55.6%;	Pred. No. 0.88;		
Matches	55;	Conservative	0;	Mismatches 44;
			Indels	0;
			Gaps	0;

Oy	98	CAGAACACCACGGTGTATCAATAGGACACGTGTCTCGGACACACCGTGGGAAAGGACACT	157
Db	411	CAGAGCAACCGACAGGTTAAATAGGGAGTGTCTCGATACACACCACTACAAATGTTTCAG	352
Oy	158	TTGTTTCTTATCACTCGACACAGCAGCTCCCAATC	196
Db	351	TTGACACTGCGCCACAGTATAGGACACCAACAGTATC	313

RESULT 5
 US-09-411-329C-6/c
 : Sequence 6, Application US/09411329C
 : Patent No. 6261820
 : GENERAL INFORMATION:
 : APPLICANT: Boone, Thomas
 : APPLICANT: Li, Huimin
 : APPLICANT: Mann, Michel
 : TITLE OF INVENTION: FIBRINOLYTICALLY ACTIVE POLYPEPTIDE
 : FILE REFERENCE: A-596
 : CURRENT APPLICATION NUMBER: US/09/411,329C
 : CURRENT FILING DATE: 1999-10-01
 : NUMBER OF SEQ ID NOS: 29
 : SOFTWARE: PatentIn version 3.0
 : SEQ ID NO 6
 : LENGTH: 609
 : TYPE: DNA
 : ORGANISM: Agkistrodon contortrix
 US-09-411-329C-6

Query Match	13.18;	Score 28.6;	DB 4;	Length 609;
Best Local Similarity	55.68;	Pred. No. 0.88;		

	Matches	Conservative	0;	Mismatches	44;	Indels	0;	Gaps
1	55							
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								
36								
37								
38								
39								
40								
41								
42								
43								
44								
45								
46								
47								
48								
49								
50								
51								
52								
53								
54								
55								
56								
57								
58								
59								
60								
61								
62								
63								
64								
65								
66								
67								
68								
69								
70								
71								
72								
73								

RESULT: b
 US-09-411-329C-19/c
 Sequence 19, Application US/09411329C
 Patent No. 6261820
 GENERAL INFORMATION:
 APPLICANT: Boone, Thomas
 APPLICANT: Li, Huimin
 APPLICANT: Mann, Michael
 TITLE OF INVENTION: FIBRINOLYTICALLY ACTIVE POLYPEPTIDE
 FILE/REFERENCE: A-596
 CURRENT APPLICATION NUMBER: US/09/411,329C
 CURRENT FILING DATE: 1999-10-01
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 19
 LENGTH: 816
 TYPE: DNA
 ORGANISM: Agkistrodon contortrix
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Fragment of fibrinolyse of Agkistrodon contortrix
 US-09-411-329C-19

Query/Match	13.1%	Score 28.6;	DB 4;	Length 816;
Best local Similarity	55.6%	Pred. No. 1;		
Matches 55; Conservative		Mismatches 44;	Indels 0;	Gaps 0

[illegible]

RESULT: 7
 US-09-411-329C-20/c
 Sequence 20, Application US/09411329C
 Patent No. 6261820
 GENERAL INFORMATION:
 APPLICANT: Boone, Thomas
 APPLICANT: Li, Huimin
 APPLICANT: Mann, Michel
 TITLE OF INVENTION: FIBRINOLYTICALLY ACTIVE POLYPEPTIDE
 FILE REFERENCE: A-596
 CURRENT APPLICATION NUMBER: US/09/411.329C
 CURRENT FILING DATE: 1999-10-01
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 20
 LENGTH: 1373
 TYPE: DNA
 ORGANISM: Agkistrodon contortrix
 FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Fragment of fibrolase of Agkistrodon contortrix
 US-09-411-329C-20

Query: match	13.1%;	Score 28.6;	DB 4;	Length 1373;
--------------	--------	-------------	-------	--------------

Db 197 CACCTGGAAGGAGCTGAGTGGCTTCATCGGAATAAGCAATGATACACAA 256
 QY 96 TCAGAACAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 155
 Db 257 CAGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 316
 QY 156 CTTTGTCTTCT 165
 Db 317 CACTGTACTCT 326

RESULT 14
 US-09-328-111-53/C
 : Sequence 53, Application US/09328111
 : Patent No. 6262333
 : GENERAL INFORMATION:
 : APPLICANT: Endege, Wilson O.
 : APPLICANT: Steinmann, Kathleen E.
 : APPLICANT: Astle, Jon H.
 : APPLICANT: Burgess, Christopher C.
 : APPLICANT: Bushnell, Steven E.
 : APPLICANT: Carroll III, Eddie
 : APPLICANT: Catino, Theodore J.
 : APPLICANT: Dertl, Adnan
 : APPLICANT: Ford, Donna M.
 : APPLICANT: Lewis, Marcia E.
 : APPLICANT: Monahan, John E.
 : APPLICANT: Schlegel, Robert
 : TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 : FILE REFERENCE: CCD-257 (US)
 : CURRENT APPLICATION NUMBER: US/09/328,111
 : CURRENT FILING DATE: 1999-06-08
 : EARLIER APPLICATION NUMBER: US 60/088,801
 : EARLIER FILING DATE: 1998-06-10
 : NUMBER OF SEQ ID NOS: 850
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 53
 : LENGTH: 632
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: (1)...(632)
 : OTHER INFORMATION: n = A,T,C or G
 : US-09-328-111-53

Query Match 12.4%; Score 27; DB 4; Length 632;
 Best Local Similarity 48.1%; Pred. No. 3.3;
 Matches 63; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 9 GCCTCATGATGCTTTTGGGCGCTTCATCAGAAATGAGAGCTGCTCAGCCCTGCA 68
 Db 575 GTCCAAATTAAAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
 QY 69 TCAGAGTTGAGTAAAGGATTAACCTCCAGAACAGCAGTGGATGAATGAGCAGTGA 128
 Db 515 GCCAAACTCCAATTCTGATTTCTCTCCNTNAGCMTCCAGCTGCTGTTANACACTTA 456
 QY 129 TCCTGACAGC 139
 Db 455 CCNTGAAGACC 445

RESULT 15
 US-08-602-716A-3
 : Sequence 3, Application US/08602716A
 : Patent No. 5962664
 : GENERAL INFORMATION:
 : APPLICANT: FRIEDHOFF, Arnold J.
 : APPLICANT: BASHAM, Daryl A.

APPLICANT: MILLER, Jeanette C.
 TITLE OF INVENTION: PSYCHOSIS PROTECTING NUCLEIC ACID,
 TITLE OF INVENTION: PEPTIDES, COMPOSITIONS AND METHODS OF USE
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROWDY AND NEIMARK
 STREET: 419 Seventh Street, N.W., Suite 300
 City: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/602,716A
 FILING DATE: 23-FEB-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/060,560
 FILING DATE: 13-MAY-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/05545
 FILING DATE: 13-MAY-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BROWDY, Roger L.
 REGISTRATION NUMBER: 25,618
 REFERENCE/DOCKET NUMBER: FRIEDHOFF-1A
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 371 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-602-716A-3

Query Match 12.3%; Score 26.8; DB 2; Length 371;
 Best Local Similarity 61.4%; Pred. No. 3;
 Matches 43; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
 QY 102 ACAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 161
 Db 185 ACCGCCATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 244
 QY 152 TTCTATGAC 171
 Db 245 ACCTATGAC 254

Search completed: April 3, 2002, 20:53:57
 Job time: 41795.sec

AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shihahara, T., Tanaka, T. and Nakamura, Y.

TITLE Direct Submission

JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: chn@elms.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES Location/Qualifiers

source 1. 2022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="COL2275"
/clone_1ib="COL"
/tissue_type="colon"
/note="cloning vector pME18SFL3"
1. 2022
/note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA"

misc_feature 1. 2022
/note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA"

BASE COUNT 612 a 472 c 453 g 485 t

ORIGIN

Query Match 100.0%; Score 154; DB 9; Length 2022;
Best Local Similarity 100.0%; Pred. No. 7.9e-38;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGATCTTTTCATTCGCTATTCAGGCTGTTGATTAAGCTGATCGAATCAGAAATATCC 60
|||||
DB 1725 ACAGATCTTTTCATTCGCTATTCAGGCTGTTGATTAAGCTGATCGAATCAGAAATATCC 1784

QY 61 AACATTGCACGAGTATCTTTGTTATTCCTCCACAGACCTCCGCCAGACAGACCTAGCTC 120
|||||
DB 1785 AACATTGCACGAGTATCTTTGTTATTCCTCCACAGACCTCCGCCAGACAGACCTAGCTC 1844

QY 121 GATGAACGCTGCTGCT 154
|||||
DB 1845 GATGAACGCTGCTGCT 1878

RESULT 2
LOCUS AF127036 2826 bp mRNA PRI 10-AUG-1999
DEFINITION Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)
ACCESSION AF127036
VERSION AF127036.1 GI:4585468
KEYWORDS mRNA, complete cds.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2826)
AUTHORS Agnel, M., Vermet, T., and Culouscou, J.-M.
TITLE Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea

JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792
REFERENCE 2 (bases 1 to 2826)
AUTHORS Agnel, M., and Culouscou, J.-M.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrieres, Rueil-Malmaison 92500, France

FEATURES Location/Qualifiers

source 1. 2826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="small intestine; colon"
1. 2826
/gene="CaCC1"
5. 2749
/gene="CaCC1"

/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 1"
/protein_id="AA025487.1"
/db_xref="GI:4585469"

BASE COUNT 875 a 623 c 632 g 696 t

ORIGIN

Query Match 100.0%; Score 154; DB 9; Length 2826;
Best Local Similarity 100.0%; Pred. No. 8.3e-38;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGATCTTTTCATTCGCTATTCAGGCTGTTGATTAAGCTGATCGAATCAGAAATATCC 60
|||||
DB 2516 ACAGATCTTTTCATTCGCTATTCAGGCTGTTGATTAAGCTGATCGAATCAGAAATATCC 2575

QY 61 AACATTGCACGAGTATCTTTGTTATTCCTCCACAGACCTCCGCCAGACAGACCTAGCTC 120
|||||
DB 2576 AACATTGCACGAGTATCTTTGTTATTCCTCCACAGACCTCCGCCAGACAGACCTAGCTC 2635

QY 121 GATGAACGCTGCTGCT 154
|||||
DB 2636 GATGAACGCTGCTGCT 2669

RESULT 3
LOCUS AX193489 3311 bp DNA PAT 15-AUG-2001
DEFINITION Sequence 1056 from Patent WO0149716.
ACCESSION AX193489
VERSION AX193489.1 GI:15211440
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3311)
AUTHORS Xu, J., Jodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J., Stolk, J.A., King, G.E., Wang, T., and Jiang, Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and methods for their use
JOURNAL CORIAX CORPORATION (US)
REFERENCE 2 (bases 1 to 3311)
AUTHORS Xu, J., Jodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J., Stolk, J.A., King, G.E., Wang, T., and Jiang, Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and methods for their use
JOURNAL CORIAX CORPORATION (US)
REFERENCE 3 (bases 1 to 3311)
AUTHORS Xu, J., Jodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J., Stolk, J.A., King, G.E., Wang, T., and Jiang, Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and methods for their use
JOURNAL CORIAX CORPORATION (US)
REFERENCE 4 (bases 1 to 3311)
AUTHORS Xu, J., Jodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J., Stolk, J.A., King, G.E., Wang, T., and Jiang, Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and methods for their use
JOURNAL CORIAX CORPORATION (US)


```

BASE COUNT      1028 a      692 c      742 g      849 t
ORIGIN
Query Match      100.0%; Score 154; DB 6; Length 3311;
Best Local Similarity 100.0%; Pred. No. 8.5e-38;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACAGATCTTTTCATTCGATTCAGGCTGTGATAGTCGATCGAATCAAGAAATATCC 60
Db 2863 ACAGATCTTTTCATTCGATTCAGGCTGTGATAGTCGATCGAATCAAGAAATATCC 2922
OY 61 AACATTCGACGAGTATCTTTGTTATTCCTCCACAGACTCCGACAGACACCTAGTCT 120
Db 2923 AACATTCGACGAGTATCTTTGTTATTCCTCCACAGACTCCGACAGACACCTAGTCT 2982
OY 121 GATGAACGCTGCTGCTCTGCTGCTTAATATTCATA 154
Db 2983 GATGAACGCTGCTGCTGCTGCTTAATATTCATA 3016

RESULT 4
AF039400      3311 bp      mRNA      PRI      14-DEC-1998
LOCUS      Homo sapiens calcium-dependent chloride channel-1 (hclcal) mRNA,
DEFINITION      complete cds.
ACCESSION      AF039400
VERSION      AF039400.1 GI:4009457
KEYWORDS
ORGANISM      human.
REFERENCE      Homo sapiens
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
      Gruber,A.D., Eblie,R.C., Ji,H.L., Schreur,K.D., Fuller,C.M. and
      Pauli,B.U.
      1 (bases 1 to 3311)
      Genomic cloning, molecular characterization, and functional
      analysis of human CLCA1, the first human member of the family of
      Ca2+-activated Cl- channel proteins
      Genomics 54 (2), 200-214 (1998)
JOURNAL      2 (bases 1 to 3311)
MEDLINE      Gruber,A.D., Eblie,R. and Pauli,B.U.
REFERENCE      Direct Submission
AUTHORS      Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
      Cornell University, Ithaca, NY 14853, USA
JOURNAL      Location/Qualifiers
FEATURES
source
1. .3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/issue_type="small intestine"
1. .3311
/gene="hclcal"
352. .3096
/gene="hclcal"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPFKSSVFLILHLLEGALSNSLTOLNNNGREGIVATIDPNP
EDTLIOQKDWYTOASLYFEATGKRFYKKNVAIILPEWTKRADYVRKLETKKA
DVLVAESTPGNDPEPTTBQNGKGEKERIHLTPDLGKRLAEGQGAFFHEMHA
LWGVDEYNNDEKFLSNGRIQAVRCSAGITGVVKKQGGSCYRKCTFKNVTGL
YKGCDFVLOSRTKASIMFAQHVDSIVFCTEONINKAPNOKNCKMLRSTWEYL
RQSEDEKKTPTPTOPNPFSLQIGORIVCLVLDKSGMAGNRLNRLNOAGOLEL
LQVELGSWGVATPDPSAAHVOSELIOINGSDRDLAKRLPAAAGSGTSGLSBSA
FYIRKKYPTDSEIVILTDGENDTISGCNFKVKGSGALIIHTVALGSAAOBLELSK
MGGLOITYASDOYONNGLDARFALSNGAVSORSIQLESKGLTLNDSMMNGTIV
DSTVGKDTLEFLITWTPQIILMDPSGOKGFGVYDKNTKMAVLIQIPGIAKGTWKY

```

```

SLQASSQTLFTVTSRASNAITPITVTSKTNKDSKRPSPLYVYANIRQASPIRLA
STALISVSNKTYTLELDNAGADAKDGYVSRYFTYDTNGRSYKRALGYN
AARRVAPDOOGALYIPGMEINDEIOMNPPPEIINKDOYHOKVCFSTSSGGSFVVS
DVNPAPIPDLPEPQOIIDIAKEIIGESLINLTWPARDDDYDGHANHYIIPISYILD
LRQFNSLOYNTTALLPKRANSEVFLRFENDITFENGDDLFATDAVADVLSKEI
SNIRVSLFIPQPPETPSDEFSAACPNIHINSTIPGHIILKIMRWIGELQLSLA
BASE COUNT      1028 a      692 c      742 g      849 t
ORIGIN
Query Match      100.0%; Score 154; DB 9; Length 3311;
Best Local Similarity 100.0%; Pred. No. 8.5e-38;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACAGATCTTTTCATTCGATTCAGGCTGTGATAGTCGATCGAATCAAGAAATATCC 60
Db 2863 ACAGATCTTTTCATTCGATTCAGGCTGTGATAGTCGATCGAATCAAGAAATATCC 2922
OY 61 AACATTCGACGAGTATCTTTGTTATTCCTCCACAGACTCCGACAGACACCTAGTCT 120
Db 2983 AACATTCGACGAGTATCTTTGTTATTCCTCCACAGACTCCGACAGACACCTAGTCT 2982
OY 121 GATGAACGCTGCTGCTCTGCTGCTTAATATTCATA 154
Db 2983 GATGAACGCTGCTGCTCTGCTGCTTAATATTCATA 3016

RESULT 5
HSJ651E10      113764 bp      DNA      PRI      27-MAY-2000
LOCUS      Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1,
DEFINITION      complete sequence.
ACCESSION      AL122002
VERSION      AL122002.16 GI:8247274
KEYWORDS      HTG.
ORGANISM      human.
REFERENCE      Homo sapiens
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
      Brown,A.
      1 (bases 1 to 113764)
JOURNAL      Direct Submission
AUTHORS      Sanger Centre, Hinxton, Cambridgeshire,
      Submitted (27-MAY-2000)
JOURNAL      Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
      CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
      requests: clonerequests@sanger.ac.uk
      On Jun 4, 2000 this sequence version replaced gi:8247030.
      During sequence assembly data is compared from overlapping clones.
      Where differences are found these are annotated as variations
      together with a note of the overlapping clone name. Note that the
      variation annotation may not be found in the sequence submission
      corresponding to the overlapping clone, as we submit sequences with
      only a small overlap as described above.
      This sequence has been finished according to sequence map criteria
      as follows. An attempt is made to resolve all sequencing problems,
      such as compressions and repeats, but not necessarily within known
      annotated human repeat sequence elements (e.g. Alu) where the
      sequence is ambiguous, there is an annotation using the 'unsure'
      feature key.
      The following abbreviations are used to associate primary accession
      numbers given in the feature table with their source databases:
      Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
      on the WORMPEP database can be found at
      http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
      was generated from part of bacterial clone contigs of human
      chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
      group. Further information can be found at
      http://www.sanger.ac.uk/HGP/Chr1
      RP4-651E10 is from the library RPC1-4 constructed at the Roswell
      Park Cancer Institute by the group of Pieter de Jong. For further
      details see http://bacpac.med.buffalo.edu/
      VECTOR: pCYPAC2
      This sequence is the entire insert of clone RP4-651E10.

```


FEATURES
Source

Location/Qualifiers
1. 113764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p22.3-31.1"
/clone="RP4-651E10"
/clone.lib="RP4-4"
/clone.lib="RP4-4" 34534 t

BASE COUNT 34890 a 21989 c 22351 g 34534 t
ORIGIN

Query Match 100.0%; Score 154; DB 9; Length 113764;
Best Local Similarity 100.0%; Pred. No. 1.4e-37;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACAGATCTTTTCATTCGCTATTCAGGCTGTGATTAAGTTCGATGAAATCAGAAATATCC 60
|||||
DB 32782 ACAGATCTTTTCATTCGCTATTCAGGCTGTGATTAAGTTCGATGAAATCAGAAATATCC 32841

OY 61 AACATTGCACGAGTATCTTTGTTATTCCTCCACACACTCCGCCAGACAGACACTAGTCC 120
|||||

DB 32842 AACATTGCACGAGTATCTTTGTTATTCCTCCACACACTCCGCCAGACAGACACTAGTCC 32901
|||||

OY 121 GATGAACGCTGCTGCTCTTCTCTTAATATTCATA 154
|||||

DB 32902 GATGAACGCTGCTGCTCTTCTCTTAATATTCATA 32935
|||||

RESULT 6
AL358950/c 164891 bp DNA HTG 23-JAN-2001
LOCUS Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN
DEFINITION
AL358950
AL358950.4 GI:12539689
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 164891)

REFERENCE

AUTHORS
TITLE
JOURNAL
Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquere@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Jan 26, 2001 this sequence version replaced gi:9988471.

COMMENT

Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquere@sanger.ac.uk
Project Information
Center project name: BA444C12
Summary Statistics
Sequencing program: XGAP4; version 4.5
Sequencing vector: plasmid: 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least Q40
Consensus quality: 163128 bases at least Q30
Insert size: 163991; sum-of-contigs
Insert size: 163373; 6.6% error; agarose-1p
Quality coverage: 5.19x in Q20 bases; sum-of-contigs Quality
coverage: 5.21x in Q20 bases; agarose-1p

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

1 34236: contig of 34236 bp in length
* 34237 34336: gap of 100 bp
* 34337 47372: contig of 13036 bp in length
* 47373 47472: gap of 100 bp
* 47473 67025: contig of 19553 bp in length
* 67026 67125: gap of 100 bp
* 67126 71239: contig of 4114 bp in length
* 71240 71339: gap of 100 bp
* 71340 83903: contig of 12564 bp in length
* 83904 84003: gap of 100 bp
* 84004 91357: contig of 7354 bp in length
* 91358 91457: gap of 100 bp
* 91458 96452: contig of 4995 bp in length
* 96453 96552: gap of 100 bp
* 96553 148189: contig of 51637 bp in length
* 148190 148289: gap of 100 bp
* 148290 159778: contig of 11489 bp in length
* 159779 159879: gap of 100 bp
* 159880 164891: contig of 5013 bp in length.

FEATURES
Source
1. 164891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-444C12"
/clone.lib="RP4-4"
/clone.lib="RP4-4" 34534 t

misc_feature

/note="assembly_fragment:00027
fragment_chain:1
clone_end:17
vector_side:left"

misc_feature

/note="assembly_fragment:00312
fragment_chain:1"

misc_feature

/note="assembly_fragment:01923
fragment_chain:1"

misc_feature

/note="assembly_fragment:02094
fragment_chain:1"

misc_feature

/note="assembly_fragment:02213
fragment_chain:1"

misc_feature

/note="assembly_fragment:01292
fragment_chain:1"

misc_feature

/note="assembly_fragment:01391
fragment_chain:1"

misc_feature

/note="assembly_fragment:02045
fragment_chain:1"

misc_feature

/note="assembly_fragment:00427
fragment_chain:1"

misc_feature

/note="assembly_fragment:00773
fragment_chain:1"

misc_feature

/note="assembly_fragment:00773
fragment_chain:1"

BASE COUNT

48381 a 32179 c 32357 g 51067 t 907 others

ORIGIN

Query Match 100.0%; Score 154; DB 2; Length 164891;
Best Local Similarity 100.0%; Pred. No. 1.5e-37;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACAGATCTTTTCATTCGCTATTCAGGCTGTGATTAAGTTCGATGAAATCAGAAATATCC 60
|||||
DB 48919 ACAGATCTTTTCATTCGCTATTCAGGCTGTGATTAAGTTCGATGAAATCAGAAATATCC 48890


```

gene      /tissue_type="intestine"
          /gene="2756"
          /gene="gob-5"
          /gene="2756"
          /gene="gob-5"
          /codon_start=1
          /protein_id="BAA33743.1"
          /db_xref="GI:3721912"
          /translation="MESLSPVFLILHLLEGVLSLILQNNNGYEGYIADHDV
          EDELIOHKDMVTOASPYLFATGKRFYFKKVALIIPESWAKPEPTPKLETFNA
          DYIVSTSPFLGNDPEPTHEIGAGGEGHILHLPDLACKKLOVOPDRPYEHAN
          FRMGVFEYNDEKFLSKGPOAVKSAITGKNVRCQGGCITNKCVIDRYG
          LYKNCVFEVDPDHONKASIMFNONINSVERCTENHNOEAPNDONCINLSTEV
          IOSEDEFKOTPTMTAOPAPATFSLQIGRIYCVLVDKSGSLNDRLNRMNASRLE
          LQTEOGSGWGVFPDASAAYOSLKOISGADRLIKHLPYVAGGTSICGLRT
          AFTVKKKPTDGGSEIVILTDGEDNITSSCPDLYVQSGAIIHTVAAGPAAKRLBOLS
          KMTGGLQRTSSQVQONCLVDPALISSGNAIAQHSIOLESRCVNLQNNNGSVT
          VDSVGGKDTFLITWTHTPTTIFIMDPSVEQNGFLDITTKVAVLYQVGTAKVGVK
          YSIOASQTLTLTVISRAASATLPITVYVKNKNGKPPSPVTVASIROGASPLR
          ASVTALIESVNGKTVLLELDNGAGADATKNDGVSRFTAFDANGRYSVKIMALGV
          TSDROAAPPKRANV IDGMIEDGEVRMNPREPRTSYVODKOLCFRTSSGSEVATN
          VPAAPIDPLFPPOCTDLDKASIOGONLNLMTAPGDYDHRASNYIIRMTSTVD
          LRDHRTSLQVNTTGLIPKESSEEFEEELGNTGNTGNTDIFALQAYDKSLKSEI
          SNARSVFIPDPEPTDSTPPCDISINSTITPEIHVAKIMKWLGEQVTLGLH"
BASE COUNT      860 a      718 c      693 g      666 t
ORIGIN

```

```

Query Match      47.8%; Score 73.6; DB 10; Length 2937;
Best Local Similarity 81.7%; Pred. No. 1.5e-12;
Matches 85; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ACAGATCTTTTCATTCGTATTCAGCGTGTGATTAAGTCGATCGAATCGAATAATCC 60
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||
Db 2526 ACAGATATCTTCTATTCATCCAGCGTGTGATTAAGTCGATCGAATCGAATAATCC 2585
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||

QY 61 AACATTCACAGAGTATCTTGTATTATTCCTCCACAGACTCCGCC 104
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||
Db 2586 AACATTCACAGAGGCTGTCTGTTCATCCCGCTCAGAGAGCCGCC 2629
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||

```

```

RESULT 11
LOCUS      AK000138      1895 bp      mRNA      PRI      22-FEB-2000
DEFINITION Homo sapiens cDNA FLJ20131 fis, clone COL06357.
ACCESSION  AK000138
VERSION     AK000138.1 GI:7020030
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (sites)
          Kawabata,A., Hiki,I.T., Kobatake,N., Inagaki,H., Ikema,Y.,
          Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Odayashi,M., Nishi,T.,
          Shihahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
          NEDO human cDNA sequencing project
          Unpublished (2000)
          2 (bases 1 to 1895)
          Sugano,S., Suzuki,Y., Ota,T., Odayashi,M., Nishi,T., Isogai,T.,
          Shihahara,T., Tanaka,T. and Nakamura,Y.
          Direct Submission
          Submitted (15-FEB-2000) to the DDBJ/EMBL/Genbank databases. Sumio
          Sugano, Institute of Medical Science, University of Tokyo, Department
          of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
          (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
          Fax:81-3-5449-5416)

```

```

COMMENT
NEDO human cDNA sequencing project. supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology

```

```

FEATURES
source      Agency)
            location/Qualifiers
            1..1895
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="COL06357"
            /clone_1ib="COL"
            /tissue_type="colon"
            /note="cloning vector pME18SFL3"
            66..1475
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="BAA90969.1"
            /db_xref="GI:7020031"
            /translation="MSKINGGSHFVYSDAQNNGILDAFGALTSNGTDLSSKSIQLES
            KGLITLSNMAMNTYVILDSVGGDTFFPLITMNSLSPSISLMDPSGTMETFTYDASK
            MAYLSTIGTAKVGTMAVNLQAKNPELTITVTSRANSSVPPITVNAKNNKDVNSFP
            SPMTIVAEILQGVAPVLANVTAFLPISONGHTEVLELDNGAGADSEKNDGVSRFT
            AYENKRSYSLKGVAPGAGANTARLKLPLPLRAVYIPGVVNGIEANPPPEIDEDTQ
            TTLEDSRTSGAGAFVVSQSLPDPQYPSQITLDATVHDKIITWTAPGND
            VGVQVYIIRISASITLDRSPDPAIDVNTTDSPEAKNSKSPAFKPEVISENATH
            IRLAISIDKSNTSKYSNIAOVTLTIPQANPDIDPTPTPTPKSHNSGVNISTL
            VLSVGSVIVNLTSTTI"
BASE COUNT      658 a      352 c      357 g      528 t
ORIGIN

```

```

Query Match      40.8%; Score 62.8; DB 9; Length 1895;
Best Local Similarity 63.0%; Pred. No. 3.4e-09;
Matches 97; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1 ACAGATCTTTTCATTCGTATTCAGCGTGTGATTAAGTCGATCGAATCGAATAATCC 60
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||
Db 1236 ACCACATATTATTTATTCATTAAGAATAGATATAAGCAATTTGACATCAAAAGTATCC 1295
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||

QY 61 AACATTCACAGAGTATCTTGTATTATTCCTCCACAGACTCCGCCAGACACTAGTCT 120
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||
Db 1296 AACATTCACAGAGTATCTTGTATTATTCCTCCACAGACTCCGTGATGATGATCTACT 1355
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||

QY 121 GATGAACGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 154
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||
Db 1356 CTTACTCTTACTCTACTCTGATTAAGTATCA 1389
    ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 |||||

```

```

RESULT 12
LOCUS      AF127035      3204 bp      mRNA      PRI      11-AUG-1999
DEFINITION Homo sapiens calcium-activated chloride channel protein 2 (CaCC2)
ACCESSION  AF127035
VERSION     AF127035.1 GI:5726288
KEYWORDS   mRNA, complete cds.
SOURCE     human.
ORGANISM   human.

```

```

REFERENCE  1 (bases 1 to 3204)
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          Agnel,M., Vernat,T. and Culoussou,J.M.
          Identification of three novel members of the calcium-dependent
          chloride channel (CaCC) family predominantly expressed in the
          digestive tract and trachea
          FEBS Lett. 455 (3), 295-301 (1999)

```

```

JOURNAL    MEDLINE
PUBMEDID   99364503
MEDLINE    10437792

```

```

REFERENCE  2 (bases 1 to 3204)
          Agnel,M. and Culoussou,J.-M.
          Direct Submission
          Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des
          Carrieres, Neuilly-Malmaison 92500, France

```

```

TITLE      location/Qualifiers
JOURNAL    1..3204
FEATURES    /organism="Homo sapiens"
            /db_xref="taxon:9606"

```



```

/note="assembly_fragment:00102
fragment_chain:1"
3393..10168
/note="assembly_fragment:00109
fragment_chain:1"
10269..21700
/note="assembly_fragment:00228
fragment_chain:2"
21801..29662
/note="assembly_fragment:00045
fragment_chain:2"
29763..31862
/note="assembly_fragment:00277
fragment_chain:3"
31963..36577
/note="assembly_fragment:00849
fragment_chain:3"
36678..43943
/note="assembly_fragment:00578
fragment_chain:4"
44044..52906
/note="assembly_fragment:00490
fragment_chain:4"
53007..65117
/note="assembly_fragment:01028
fragment_chain:5"
65218..70699
/note="assembly_fragment:00015
fragment_chain:5"
70800..82141
/note="assembly_fragment:00071"
82242..87387
/note="assembly_fragment:00242"
87488..91496
/note="assembly_fragment:00323"
91597..104147
/note="assembly_fragment:00430"
104248..107008
/note="assembly_fragment:00435"
107109..110463
/note="assembly_fragment:00756"
110564..112670
/note="assembly_fragment:00793"
112771..116165
/note="assembly_fragment:00810"
116266..118389
/note="assembly_fragment:00912"
118490..129046
/note="assembly_fragment:01001"
129147..136534
/note="assembly_fragment:01002"
136635..140718
/note="assembly_fragment:01014"
140819..144918
/note="assembly_fragment:01015"
145019..149118
/note="assembly_fragment:01016"
149219..153318
/note="assembly_fragment:01017"
153419..157518
/note="assembly_fragment:01018"
157619..161718
/note="assembly_fragment:01019"
161819..165918
/note="assembly_fragment:01020"
166019..170118
/note="assembly_fragment:01021"
170219..174318
/note="assembly_fragment:01022"
174419..178518
/note="assembly_fragment:01023"
178619..182718
/note="assembly_fragment:01024"
182819..186918
/note="assembly_fragment:01025"
187019..191118
/note="assembly_fragment:01026"
191219..195318
/note="assembly_fragment:01027"
195419..199518
/note="assembly_fragment:01028"
199619..203718
/note="assembly_fragment:01029"
203819..207918
/note="assembly_fragment:01030"
208019..212118
/note="assembly_fragment:01031"
212219..216318
/note="assembly_fragment:01032"
216419..220518
/note="assembly_fragment:01033"
220619..224718
/note="assembly_fragment:01034"
224819..228918
/note="assembly_fragment:01035"
229019..233118
/note="assembly_fragment:01036"
233219..237318
/note="assembly_fragment:01037"
237419..241518
/note="assembly_fragment:01038"
241619..245718
/note="assembly_fragment:01039"
245819..249918
/note="assembly_fragment:01040"
250019..254118
/note="assembly_fragment:01041"
254219..258318
/note="assembly_fragment:01042"
258419..262518
/note="assembly_fragment:01043"
262619..266718
/note="assembly_fragment:01044"
266819..270918
/note="assembly_fragment:01045"
271019..275118
/note="assembly_fragment:01046"
275219..279318
/note="assembly_fragment:01047"
279419..283518
/note="assembly_fragment:01048"
283619..287718
/note="assembly_fragment:01049"
287819..291918
/note="assembly_fragment:01050"
292019..296118
/note="assembly_fragment:01051"
296219..300318
/note="assembly_fragment:01052"
300419..304518
/note="assembly_fragment:01053"
304619..308718
/note="assembly_fragment:01054"
308819..312918
/note="assembly_fragment:01055"
313019..317118
/note="assembly_fragment:01056"
317219..321318
/note="assembly_fragment:01057"
321419..325518
/note="assembly_fragment:01058"
325619..329718
/note="assembly_fragment:01059"
329819..333918
/note="assembly_fragment:01060"
334019..338118
/note="assembly_fragment:01061"
338219..342318
/note="assembly_fragment:01062"
342419..346518
/note="assembly_fragment:01063"
346619..350718
/note="assembly_fragment:01064"
350819..354918
/note="assembly_fragment:01065"
355019..359118
/note="assembly_fragment:01066"
359219..363318
/note="assembly_fragment:01067"
363419..367518
/note="assembly_fragment:01068"
367619..371718
/note="assembly_fragment:01069"
371819..375918
/note="assembly_fragment:01070"
376019..380118
/note="assembly_fragment:01071"
380219..384318
/note="assembly_fragment:01072"
384419..388518
/note="assembly_fragment:01073"
388619..392718
/note="assembly_fragment:01074"
392819..396918
/note="assembly_fragment:01075"
397019..401118
/note="assembly_fragment:01076"
401219..405318
/note="assembly_fragment:01077"
405419..409518
/note="assembly_fragment:01078"
409619..413718
/note="assembly_fragment:01079"
413819..417918
/note="assembly_fragment:01080"
418019..422118
/note="assembly_fragment:01081"
422219..426318
/note="assembly_fragment:01082"
426419..430518
/note="assembly_fragment:01083"
430619..434718
/note="assembly_fragment:01084"
434819..438918
/note="assembly_fragment:01085"
439019..443118
/note="assembly_fragment:01086"
443219..447318
/note="assembly_fragment:01087"
447419..451518
/note="assembly_fragment:01088"
451619..455718
/note="assembly_fragment:01089"
455819..459918
/note="assembly_fragment:01090"
460019..464118
/note="assembly_fragment:01091"
464219..468318
/note="assembly_fragment:01092"
468419..472518
/note="assembly_fragment:01093"
472619..476718
/note="assembly_fragment:01094"
476819..480918
/note="assembly_fragment:01095"
481019..485118
/note="assembly_fragment:01096"
485219..489318
/note="assembly_fragment:01097"
489419..493518
/note="assembly_fragment:01098"
493619..497718
/note="assembly_fragment:01099"
497819..501918
/note="assembly_fragment:01100"
502019..506118
/note="assembly_fragment:01101"
506219..510318
/note="assembly_fragment:01102"
510419..514518
/note="assembly_fragment:01103"
514619..518718
/note="assembly_fragment:01104"
518819..522918
/note="assembly_fragment:01105"
523019..527118
/note="assembly_fragment:01106"
527219..531318
/note="assembly_fragment:01107"
531419..535518
/note="assembly_fragment:01108"
535619..539718
/note="assembly_fragment:01109"
539819..543918
/note="assembly_fragment:01110"
544019..548118
/note="assembly_fragment:01111"
548219..552318
/note="assembly_fragment:01112"
552419..556518
/note="assembly_fragment:01113"
556619..560718
/note="assembly_fragment:01114"
560819..564918
/note="assembly_fragment:01115"
565019..569118
/note="assembly_fragment:01116"
569219..573318
/note="assembly_fragment:01117"
573419..577518
/note="assembly_fragment:01118"
577619..581718
/note="assembly_fragment:01119"
581819..585918
/note="assembly_fragment:01120"
586019..590118
/note="assembly_fragment:01121"
590219..594318
/note="assembly_fragment:01122"
594419..598518
/note="assembly_fragment:01123"
598619..602718
/note="assembly_fragment:01124"
602819..606918
/note="assembly_fragment:01125"
607019..611118
/note="assembly_fragment:01126"
611219..615318
/note="assembly_fragment:01127"
615419..619518
/note="assembly_fragment:01128"
619619..623718
/note="assembly_fragment:01129"
623819..627918
/note="assembly_fragment:01130"
628019..632118
/note="assembly_fragment:01131"
632219..636318
/note="assembly_fragment:01132"
636419..640518
/note="assembly_fragment:01133"
640619..644718
/note="assembly_fragment:01134"
644819..648918
/note="assembly_fragment:01135"
649019..653118
/note="assembly_fragment:01136"
653219..657318
/note="assembly_fragment:01137"
657419..661518
/note="assembly_fragment:01138"
661619..665718
/note="assembly_fragment:01139"
665819..669918
/note="assembly_fragment:01140"
670019..674118
/note="assembly_fragment:01141"
674219..678318
/note="assembly_fragment:01142"
678419..682518
/note="assembly_fragment:01143"
682619
```

Query Match	Best Local Similarity	40.8%	Score 62.8	DB 2	length 140718
Matches	97	Conservative	0	Mismatches	57
QY	1	ACAGATCTTTTCATTCGATTCAGGCTGTGTATGAGTGCATCTGAAATCAGAAATATCC	60		
Db	40754	ACCCAGCATTTTATTTGGCATTTAAAGSTATAGATTTAAAGCAATTTGACATCAAAAGATGTC	40695		
QY	61	AACATTCGACAGATATCTTTTGTATTTCTCCACAGACTCGCCAGACACCTAGTCT	120		
Db	40694	AACATTCGACAGATATCTTTTGTATTTCTCCACAGAAATCTCGATGACATTTGATCTACT	40635		
QY	121	GATGAAGCTCGTCCTTCCTTAATATTCATA	154		
Db	40634	CTTACTCTACTCTACTCTGATTAAGATCTATA	40601		

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES						
14	AK000072	ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.	AK000072	AK000072.1	GI:7019922	oigo:capping; fis (full insert sequence).	Homo sapiens colon CDNA to mRNA, clone_1lib:COL clone:COL01613.	1 (sites)	Kawabata,A., Hikiji,T., Kobayake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shihabara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.	NEBO human CDNA sequencing project	Unpublished (2000)	2 (bases 1 to 3221)	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shihabara,T., Tanaka,T. and Nakamura,Y.						
15	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
16	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
17	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
18	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
19	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
20	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
21	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
22	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
23	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
24	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
25	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;	Query Match	1.3221	Location/Qualifiers	/organism="Homo sapiens"	/db_xref="taxon:9606"	/clone="COL01613"	/clone_lib="COL"	/lssue_type="colon"	/note="cloning vector PME18SFL3"	1.3221	/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"	CHANNEL PROTEIN"	CHANEL PROTEIN"
26	AX092338	Sequence 69 from Patent WO0116318.	AX092338	AX092338.1	GI:134444483	36.8%; Score 56.6; DB 9; Length 3221; Best Local Similarity 65.6%; Pred. No.3.2e-07; Matches 99; Conservative 0; Mism													

KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 3265)
Eaton,D.L., Filvarolf,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
Wood,W.I.

TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same

JOURNAL Patent: WO 0116318-A 69 08-MAR-2001;
Genentech, Inc. (US)

FEATURES
Source location/Qualifiers
1..3265
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1159 a 596 c 632 g 878 t
ORIGIN

Query Match 36.8%; Score 56.6; DB 6; Length 3265;
Best Local Similarity 65.6%; Pred. No. 3.2e-07;
Matches 99; Conservative 0; Mismatches 49; Indels 3; Gaps 1;

OY 1 ACAGATCTTTTCATGCTATTCAGCGCTGTTGATAGGTGATGAAATCAGAAATATGCC 60
DB 2539 ACCCACATATTATTGCGCATTAAGATATAGATTAAGCAATTTGACATCAAAAGTATCC 2598
OY 61 AACATTGCACAGATATCTTTGTTATTCCTCCACAGACTCCGCCAGAGA--CACCTAGT 117
DB 2599 AACATTGCACAAAGTACTTTGTTATTCCTCCACAAATCCGTGATGACATTTGATCCTACA 2658
OY 118 CCTGATGAAAGCTGCTGCTTGTCTCTAATA 148
DB 2659 CCTACTCTACTCTCTACTCTCTACTCTCTGATA 2689

Search completed: April 4, 2002, 03:05:40
Job time: 60622 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:41 ; Search time 16681 Seconds

(without alignments)
159,760 Million cell updates/sec

Title: US-09-049-696-14

Perfect score: 248
Sequence: 1 ACCTGACGCGCAATTCAC.....TTGAAATGCGACAGATCTT 248

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gp_est1:*
11: gp_est2:*
12: gp_hic:*
13: gp_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	248	100.0	734	10	AI660957 w120408.x
2	247	99.6	561	10	AI802756 w18a04.x
3	237	95.6	310	10	AA297150 EST112734
4	236	95.2	598	11	BE927204
5	233.4	94.1	712	10	AM009763 w87b04.x
6	219	88.3	504	10	AI802693 w117c01.x
7	214.4	86.5	501	10	AI721275 as82h08.x
8	213.4	86.0	716	10	AI660234 w68g02.x
9	171	69.0	528	11	BF083270 RC1-CT024
10	156.4	63.1	466	11	BG195179 RST14358
11	149.4	60.2	507	11	AA623217 v114d07.r
12	149.4	60.2	528	11	BG965506 602830872

13	149.4	60.2	554	10	AA623196 v114d07.r
14	149.4	60.2	629	10	AA518112 v123c06.r
15	149.4	60.2	646	10	AA691335 v514h06.r
16	149.4	60.2	732	11	BG963719 602831354
17	149.4	60.2	812	11	BG968878 602836186
18	149.4	60.2	2933	12	AK008659 Mus muscu
19	147.8	59.6	2915	12	AK007466 Mus muscu
20	143	57.3	552	10	AA245498 my51e11.r
21	142	57.3	552	10	AA688953 v504h07.r
22	129	52.0	687	11	BG286238 602383065
23	127.4	51.4	813	11	BG216984 RST6685
24	125.8	50.7	370	10	AA711143 v156d09.r
25	125.8	50.7	488	10	AA682395 RC0-CT038
26	123.4	49.8	301	10	AA689856 v162c11.r
27	119.4	48.1	441	11	W41083 mc41c0.r1
28	118.4	47.7	699	10	AI747448 u115c01.x
29	111	44.8	555	11	BI291875 UT-R-DNO-
30	106.6	43.0	297	10	AA829159 v958c07.r
31	105.4	42.5	315	11	BF083396 RC6-CT043
32	101	40.7	794	11	BG220001 RST39773
33	100.4	40.5	378	10	AA296955 EST112726
34	97.4	39.3	482	10	AA711228 v70c11.r
35	94.4	38.1	452	10	AA822200 v936d05.r
36	94.2	38.0	527	10	BE137034 u958h06.y
37	91.8	37.0	611	10	BE19720 u952c02.x
38	91	36.7	479	13	AZ827199 ZM0103E04
39	88.8	35.8	243	11	W20635 mc16f08.r1
40	88.6	35.7	488	10	AA733415 v755b06.r
41	85	34.3	455	11	BF083321 RC1-CT024
42	83	33.5	512	11	BI291975 UT-R-DNO-
43	79	31.9	392	10	BE374865 601226b5
44	79	31.9	676	11	BI105924 602892170
45	79	31.9	824	11	BI109423 602897324

ALIGNMENTS

RESULT 4 1
LOCUS AI660957/c
DEFINITION w120408.x1 Soares-Dieckgraefe_colon_NHUC Homo sapiens cDNA clone
IMAGE:2351151 3' similar to TR:088826 O88826 GDB-5 PROTEIN. ; mRNA
sequence.
ACCESSION AI660957
VERSION AI660957.1 GI:4764540
KEYWORDS EST.
SOURCE human,
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 734)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 820 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 477.
FEATURES
Location/Qualifiers
1..734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2351151"
/clone_lhb="Soares-Dieckgraefe_colon_NHUC"
/tissue_type="colonic mucosa from 5 ulcerative colitis
patients"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a

patients" = "DH10B (phage-resistant)." /lab_host="Organ: colon; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-GTTTCACATCTGAGATGGAGCCGCCGCTTATTTTTTTTTTTT 3'] , double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckraete (Washington University, dieckreim.wustl.edu); colonic mucosa represents a range of disease involvement from mild cryptitis to severe ulceration, fibrosis, and degeneration. Library constructed by Bento Soares and M. Fatima Bonaldo. "

Query	Match	99.6%	Score 247	DB 10	Length 561
Best local	Similarity 99.6%	Pred. No. 3e-59	1	Indels 0	Gaps 0
Matches 247	Conservative	0	Mismatches	1	Indels 0
QY	1	ACCTGAGGCGGAAATTCACGSGGGGACGTCTCATTTAATCTGACTTGGACACGCTCTCGGG	60		
Es	536	ACCTGAGGCGGAAATTCACGSGGGGACGTCTCATTTAATCTGACTTGGACACGCTCTCGGG	477		
QY	51	ATGATTATGACCATGGAACAGCTCACAAAGTATATCATTTGGAATTAAGTACAAAGTATTTCTGG	120		
Db	476	ATGATTATGACCATGGAACAGCTCACAAAGTATATCATTTGGAATTAAGTACAAAGTATTTCTGG	417		
QY	121	ATCTCAGACGACAGTTCATGTAATCTCTCAAGTGAATACGATACGCTCTCATCCCAAGG	180		
Db	416	ATCTCAGACGACAGTTCATGTAATCTCTCAAGTGAATACGATACGCTCTCATCCCAAGG	357		
QY	181	AAGCCAACTCTGAGAGAAGTCTTTTGTGTTTAAACGAAACAATCTTCTTGGAAATGGCA	240		
Db	346	AAGCCAACTCTGAGAGAAGTCTTTTGTGTTTAAACGAAACAATCTTCTTGGAAATGGCA	297		
QY	241	CAGATCTT 248			

DB 296 CAGACCTT 289
 RESULT 3
 LOCUS AA297150 310 bp mRNA EST 18-APR-1997
 DEFINITION ES1112734 Colon I Homo sapiens CDNA 5' end, mRNA sequence.
 ACCESSION AA297150
 VERSION AA297150.1 GI:1949524
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 310)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 C., Lee,N.H., Kirkness,E.F., Weissbrock,K.G., Gocayne,J.D., White
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
 L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S., Glodke,A.,
 Guelm,C.L., Hannu,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palauques,R.F., McDonald,J.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,L.J., Saudek,D.M., Shibley,R.,
 Small,K.V., Spraggs,T.A., Ulteback,T.R., Weidman,J.F., Li,Y.,
 Dembekar,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Binkley,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A., He,M.W.
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozaik,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannom
 M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT Other ESTs: TNC167738
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M3 Reverse.

FEATURES
Source Location/Qualifiers
1..310
/organism="Homo sapiens"
/db_xref="ATCC (Inhost):127942"
/db_xref="taxon:9606"
/clone_lib="Colon I"
/dev_stage="adult"
/note="Organ: colon; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 97 a 61 c 58 g 91 t 3 others
ORIGIN

Query Match 95.6%; Score 237; DB 10; Length 310;
Best Local Similarity 99.2%; Pred. No. 1.9e-56;
Matches 237; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 CGGAATTCACGGGGGAGCTCTATTAATCTGACCTGGACAGCTCCGGGAGATTAATG 69
DB 1 CGGAATTCACGGGGGAGCTCTATTAATCTGACCTGGACAGCTCCGGGAGATTAAT 60
QY 70 ACCATGGAACAGCTCAAGATATATCATTCGATAAGTACAGATATCTTGTATCTAGAG 129
DB 61 ACCATGGAACAGCTCAAGATATATCATTCGATAAGTACAGATATCTTGTATCTAGAG 120
QY 130 ACAAGTCAATGAATCTCTCAAGTATATCTCTCTCATCCCAAGAGCAACT 189
DB 121 ACAAGTCAATGAATCTCTCAAGTATATCTCTCTCATCCCAAGAGCAACT 180
QY 190 CTGAGGAAGCTCTTTGTTTAAACAGAAACATCTTTGAAATGCGACAGATCTT 248
DB 181 CTGAGGAAGCTCTTTGTTTAAACAGAAACATCTTTGAAATGCGACAGATCTT 239

RESULT 4
LOCUS BE927204 598 bp mRNA EST 02-OCT-2000
DEFINITION RCI-CT0249-240800-211-h10 CT0249 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE927204
VERSION BE927204.1 GI:10453280
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 598)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,
Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,
Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H.,
Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare,
M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and
Simpson, A. J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE
COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: 55-11-2704922
Fax: 55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?rl=62-RC1-CT0249-240>)
800-211-h10ct3-2000-08-24(64-1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 598.

FEATURES
Source Location/Qualifiers
1..598
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0249"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTS PCR (U.S. Letters Patent application No. 196
/16 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 150 a 133 c 130 g 185 t
ORIGIN

Query Match 95.2%; Score 236; DB 11; Length 598;
Best Local Similarity 99.6%; Pred. No. 3.7e-56;
Matches 247; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACTGTAAGGGGAAATTCACGGGGGAGCTCTATTAATCTGACCTGGAGCTGGGG 60
DB 301 ACTGTAAGGGGAAATTCACGGGGGAGCTCTATTAATCTGACCTGGAGCTGGGG 242
QY 31 ANGATTAAGCAATGGAACAGCTCAAGATATCATTCGATAAGTACAGATATCTT 120
DB 241 ATGATTAAGCAATGGAACAGCTCAAGATATCATTCGATAAGTACAGATATCTT 182
QY 121 ATCTCAGAGCAAGTCAATGAATCTCTTCAAGTGAATCTACTGCTCATCCAAAG 180
DB 181 ATCTCAGAGCAAGTCAATGAATCTCTTCAAGTGAATCTACTGCTCATCCAAAG 122
QY 181 AACCAACTCTGAGGAAGTCTTTTGTAAACAGAAACATCTTTGAAATGCGA 240
DB 121 AACCAACTCTGAGGAAGTCTTTTGTAAACAGAAACATCTTTGAAATGCGA 63

QY 211 CAGATCTT 248
DB 212 CAGATCTT 55

RESULT 5
LOCUS AW009763 712 bp mRNA EST 08-MAR-2000
DEFINITION WS87b04.x1 NCI-CGAP_C03 Homo sapiens cDNA IMAGE:2504911 3'
ACCESSION AW009763
VERSION AW009763.1 GI:5858541
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 712)
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL MEDLINE

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/btrp/image/image.html
Insert length: 807 Std Error: 0.00
Seq primer: -400P from Glibco
High quality sequence stop: 451.
Location/Qualifiers

FEATURES

source
1. 712
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2504911"
/clone_lib="NCI-CGAP_C03"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization."
BASE COUNT 217 a 130 c 143 g 222 t
ORIGIN

Query Match 94.1%; Score 233.4; DB 10; Length 712;
Best Local Similarity 97.5%; Pred. No. 2e-55;
Matches 237; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

6 AAGGGGAATTCACGGGGGAGCTCATTTATCTGACCTGAGCAGCTCCGGGATGAT 65
|||||
662 AGGGGTCATCTACGAGGGGCGAGTCTCATTTATCTGACCTGAGCAGCTCCGGGATGAT 623
66 TATGACATGGAACAGCTCACAAGTATATCATTCGATTAAGTACAAGTATCTGATCTC 125
622 TATGACATGGAACAGCTCACAAGTATATCATTCGATTAAGTACAAGTATCTGATCTC 563
126 AGAGCAAGTAAAGAACTCTTCAAGTACTACTGCTGCTGATCCCAAGGAAGCC 185
562 AGAGCAAGTAAAGAACTCTTCAAGTACTACTGCTGCTGATCCCAAGGAAGCC 503
186 AACTCTGAGGAGCTCTTTTGTAAACGAGAAACATTTCTTGAATAATGGCAGAT 245
|||||
502 AACTCTGAGGAGCTCTTTTGTAAACGAGAAACATTTCTTGAATAATGGCAGAT 443
246 CTT 248
|||||
442 CTT 440

RESULT 6
A1802693/c 504 bp mRNA EST 18-DEC-1999
LOCUS A1802693 504 bp mRNA EST 18-DEC-1999
DEFINITION w17c01.x1 Soares-Dieckgraefe_colon_NHUC Homo sapiens cDNA clone
IMAGE:2350848 3' similar to TR:088826 088826 GDB-5 PROTEIN.; mRNA
sequence.
ACCESSION A1802693
VERSION A1802693.1 GI:5368165
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 504)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 727 Std Error: 0.00
Seq primer: -400P from Glibco
High quality sequence stop: 392.
Location/Qualifiers

FEATURES

source
1. 504
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2350848"
/clone_lib="Soares-Dieckgraefe_colon_NHUC"
/tissue_type="colonic mucosa from 5 ulcerative colitis patients"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGGCGGCGCTGCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieck@im.wustl.edu); colonic mucosa represents a range of disease involvement from mild cryptitis to severe ulceration, fibrosis, and degeneration. Library constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 150 a 85 c 103 g 163 t 3 others
ORIGIN

Query Match 88.3%; Score 219; DB 10; Length 504;
Best Local Similarity 98.6%; Pred. No. 2.2e-51;
Matches 219; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

37 AGTCTATTATCTGACTGGACAGCTCTCGGGATGATTTATGACATGGAACAGCTCAC 86
504 AGTCTATTATCTGACTGGACAGCTCTCGGGATGATTTATGACATGGAACAGCTCAC 445
57 AAGTATATCTGATTAAGTACAAGTATCTTGTATCTGACAGCAAGTATCAATGAATCT 146
444 AAGTATATCTGATTAAGTACAAGTATCTTGTATCTGACAGCAAGTATCAATGAATCT 385
147 CTTCAAGTAACTAGCTGCTCTCATCCCAAGCAAGCAACTGAGAGAGCTTTTG 206
344 CTTCAAGTAACTAGCTGCTCTCATCCCAAGCAAGCAACTGAGAGAGCTTTTG 325
217 TTTAAGCAAGAAACATTTGAAATGACACAGATCT 248
|||||
324 TTTAAGCAAGAAACATTTGAAATGACACAGATCT 283

RESULT 7
A1721275/c 501 bp mRNA EST 10-JUN-1999
LOCUS A1721275 501 bp mRNA EST 10-JUN-1999
DEFINITION as62h08.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
IMAGE:235263 3' similar to SW:BCLC_BOVIN P54281 EPTITHELIAL
CHLORIDE CHANNEL PROTEIN.; mRNA sequence.
ACCESSION A1721275
VERSION A1721275.1 GI:5036531
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 501)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,

TITLE	JOURNAL	MEDLINE	COMMENT
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A.	97 (7), 3491-3496 (2000)	
Contact: Simpson A.J.G.		20202863	

FEATURES	Location/Qualifiers
source	1. .528

RESULT	10
BGI95179	
LOCUS	BGI95179 466 bp mRNA
DEFINITION	RSTI4358 Athensys RAGE Library Homo sapiens CDNA, mRNA sequence
ACCESSION	BGI95179
VERSION	BGI95179.1 GI:13716866
KEYWORDS	EST.
SOURCE	human.

and Ducra, M.
Creation of Genome-wide Protein Expression Libraries using Random
Activation of Gene Expression
Nat. Biotechnol. 19 (5), 440 (2001) In press
COMMENT: JOURNAL
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 466.
FEATURES
Location/Qualifiers
Source
1..466
"micro_coding"

BASE COUNT	155 a	92 c	74 g	144 t	1 others
ORIGIN					

RESULT	11
AA623217	
LOCUS	507 bp mRNA EST 14-OCT-1997
DEFINITION	Vll4d07.i1 Barsheed mouse proximal colon MRLRB6 Mus musculus cDNA
DESCRIPTION	clone IMAGE:503757 5' similar to TR.G1184066 G1184066 CALCIUM-ACTIVATED CHLORIDE CHANNEL. ; mRNA sequence.
ACCESSION	AA623217
VERSION	AA623217.1 GI:2527093
KEYWORDS	EST.
SOURCE	house.mouse.

us Page Blank (uspto)

Human colon specific genes and their expression products - detection of which, in non-colon tissue samples, can be used as indication of colon cancer metastasis

Claim 1: Fig 5; 60pp; English.

13 CDNA clones (AA145880-92), most of them partial clones, correspond to human colon specific genes, designated CSG1, CSG2, etc., that are primarily expressed in tissues derived from the colon. CSG7 and CSG10 show reduced expression in colon cancer cells as compared to that in normal cells; the remaining genes are overexpressed in colon cancer. The partial cDNA sequences can be used to isolate full-length clones and genomic clones including the complete gene. CSG nucleic acids can be used to produce CSG polypeptides (see also AA06545-53) in transformed host cells, as probes to detect disorders of the colon, partic. colon cancer and colon cancer metastasis, and in gene therapy.

Sequence 878 BP; 257 A; 179 C; 188 G; 241 T; 13 other;

Query Match 100.0%; Score 248; DB 18; Length 878;

Best Local Similarity 100.0%; Pred. No. 3.2e-68; Mismatches 0; Indels 0; Gaps 0;

Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCTGAAGCGGAAATTCACGGGGGCACTCATTAATCTGACCTGGACGCTCGGGG 60
|||||
DB 306 acctgaagcggaattcaccgagctcattcaatcgtgacgtccctgagg 365
61 ATGATTATGACCATGAGACAGCTCACAGTATATCATTTGAATAGATATTTCTTG 120
|||||
DB 366 atgattatgaccatggaacagctcacagatatacttcgataaagatattcttg 425
QY 121 ATCTCAGACAGCAATTCATGAATCTCTCAAGTAAATCTGCTGTCATCCCAAG 180
|||||
DB 426 atctcagagcaagcttcaatgaatctctcaagtaactcgtctcatcccaagg 485
QY 181 AAGCCAACTCTGAGGAGCTTTTGTTTAAACCAAGAAACATTACTTTGAAATGCA 240
|||||
DB 486 aagccaaactctgaggaagctcttctgttaaacaccagaaacattcttgaagaatggca 545
QY 241 CAGATCTT 248
|||||
DB 546 cagatctt 553

RESULT 2

AAV16672 ID AAV16672 standard; cDNA; 878 BP.

AC AAV16672;

XX 22-JUN-1998 (first entry)

XX Polynucleotide sequence of a colon-specific gene.

XX Colon-specific gene; probe; detection; expression; human;

XX diagnostic assay; colon cancer; antibody; screening; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Key 2..685

XX Key CDS

XX Key US5733748-A.

XX Key 31-MAR-1998.

XX Key 06-JUN-1995; 95US-0469667.

XX Key 06-JUN-1995; 95US-0469667.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen C, Yu G;

XX WPI: 1998-229823/20.

XX E-PSDB; AAM46879.

XX Colon-specific nucleic acids - useful as probes for detecting colon cancer micrometastases

XX Claim 15; Fig 5A-B; 51pp; English.

AA16668-81 represent polynucleotide sequences of partial or full-length cDNA clones of colon-specific genes. The polynucleotides can be used as probes to detect expression of the corresponding human genes, e.g. in diagnostic assays for detecting micrometastases of colon cancer. Recombinant cells containing the polynucleotides can be used to produce the polypeptides, in order that antibodies can be raised and used in further screening or diagnostics.

Sequence 878 BP; 257 A; 179 C; 188 G; 241 T; 13 other;

Query Match 100.0%; Score 248; DB 19; Length 878;

Best Local Similarity 100.0%; Pred. No. 3.2e-68; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCTGAAGCGGAAATTCACGGGGGCACTCATTAATCTGACCTGGACGCTCGGGG 60
|||||
DB 306 acctgaagcggaattcaccgagctcattcaatcgtgacgtccctgagg 365
61 ATGATTATGACCATGAGACAGCTCACAGTATATCATTTGAATAGATATTTCTTG 120
|||||
DB 366 atgattatgaccatggaacagctcacagatatacttcgataaagatattcttg 425
QY 121 ATCTCAGACAGCAATTCATGAATCTCTCAAGTAAATCTGCTGTCATCCCAAG 180
|||||
DB 426 atctcagagcaagcttcaatgaatctctcaagtaactcgtctcatcccaagg 485
QY 181 AAGCCAACTCTGAGGAGCTTTTGTTTAAACCAAGAAACATTACTTTGAAATGCA 240
|||||
DB 486 aagccaaactctgaggaagctcttctgttaaacaccagaaacattcttgaagaatggca 545
QY 241 CAGATCTT 248
|||||
DB 546 cagatctt 553

RESULT 3

AAH46102 ID AAV46102 standard; DNA; 2742 BP.

XX AAV46102;

XX 11-SEP-2001 (first entry)

XX Human C1CA1 coding sequence, SEQ ID NO:4.

XX Human C1CA1; goblet cell; mouse Gob-5 orthologue; drug screening;

XX expression inhibition; antisense therapy; gene therapy;

XX chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;

XX Homo sapiens.

XX Key Location/Qualifiers

XX Key 1..2742

XX Key CDS

XX Key /tag- a

XX Key /product- "Human C1CA1"

XX Key /note- "No stop codon given in the specification"

PN W0200138530-A1.
 PD 31-MAY-2001.
 XX
 XX
 PF 22-NOV-2000; 2000WO-JP08232.
 XX
 XX 24-NOV-1999; 9901P-0333479.
 PR 27-APR-2000; 2000UP-0127589.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Nakanishi A, Morita S;
 XX
 DR WPI; 2001-355935/37.
 DR P-PSDB; AAB73716.
 XX
 XX
 PT New antisense nucleotide, useful for treatment and prevention of
 PT bronchial asthma and chronic obstructive pulmonary disease -
 XX
 XX
 PS Claim 3; Page 82-84; 104pp; Japanese.
 CC
 CC The invention relates to an antisense nucleotide targeted to the mouse
 CC Gob-5 gene (coding sequence shown in AHA46101) or its human counterpart,
 CC the CLCA1 gene (coding sequence shown in AHA46102). The invention also
 CC relates to an antibody specific for the Gob-5 protein, medical and
 CC diagnostic compositions containing the antisense nucleotide or the
 CC antibody, and methods and kits for screening for compounds which inhibit
 CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 CC The antisense oligonucleotides and antibody are therefore useful for the
 CC treatment and prevention of bronchial asthma and chronic obstructive
 CC pulmonary disease. The present sequence represents the human CLCA1
 CC gene coding sequence.
 CC
 SO Sequence 2742 BP; 833 A; 616 C; 623 G; 670 T; 0 other;

Query Match 100.0%; Score 248; DB 22; Length 2742;
 Best Local Similarity 100.0%; Pred. No. 4.8e-68;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCTGAAGCGGGAATTTCACGGGGGAGCTCATTAATCTGATGGACAGCTCTGGGG 60
 DB 2273 accggaagcggaatttcacgggggagctctcatlaactgagcttgagcagctccg 2332
 QY 61 ATGATTATGACCATGGAACAGCTCACAAGATATATCATTTGATTAAGTACGATTTCTTG 120
 DB 2333 atgattatgacctggaacagctcacaagatatactatcgtgaataagacagatattctg 2392
 QY 121 ATCTCAGAGACAAGTTCAATGAATCTTCAAGTGAATTAAGTCTGCTCTGATCCCAAGG 180
 DB 2393 atctcagagacaagttcaatgaatcttcaagtgaaatactactgctctcaccacaag 2452
 QY 181 AAGCCCACTCTGAGGAAGCTTTTGTTAACGAGAAACATTACTTTGAAAATGCGCA 240
 DB 2453 aagcccaactctgaggaagcttttgttaacgagaaacattacttltgaaaatgagca 2512
 QY 241 CAGATCTT 248
 DB 2513 cagatctt 2520

RESULT 4
 AAF81927
 ID AAF81927 standard; cDNA; 2745 BP.
 XX
 XX AAF81927;
 AC
 XX 13-JUN-2001 (first entry)
 DT
 XX Human ICACC-1 nucleotide sequence.
 DE
 XX ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
 KM Interleukin 9 induced calcium activated chloride channel; IL-9;

KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
 KW anti-inflammatory; immunomodulatory; cystic fibrosis;
 KW inflammatory bowel disease; autoimmune disease; ss.
 XX
 XX Homo sapiens:
 XX
 XX Key 1.2745
 XX CDS /tag- a
 XX /product- "ICACC-1"
 XX /note- "IL-9 induced calcium activated chloride channel"
 XX
 XX W0200138530-A1
 XX
 XX 10-SEP-1999.
 XX
 XX 03-MAR-1999; 99WO-US04703.
 XX
 XX 03-MAR-1998; 98US-0076815.
 XX
 XX (MAGNIN PHARM INC.
 XX
 XX Hoeyd KJ, Levitt RC, Maloy WL, Louhede J, McLane M;
 XX Nicolaidis NC, Zhou Y, Dong Q;
 XX WPI; 1999-550979/46.
 XX R-PSDB; AAB74824.
 XX
 XX New nucleic acid encoding calcium activated chloride channel, used to
 XX identify, e.g. specific modulators for treating atopic allergy -
 XX
 XX
 PS Claim 1; Fig 4B; 75pp; English.
 XX
 XX The present sequence encodes the human interleukin 9 (IL-9) induced
 XX calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
 XX have anti-allergic, anti-asthmatic, anti-inflammatory and
 XX immunomodulatory activities. Compounds (A) that downregulate ICACC are
 XX used to alleviate asthma (or more generally atopic allergy), while those
 XX (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
 XX inflammatory bowel disease (IBD) (or other autoimmune diseases).
 XX Measuring levels of ICACC-1 can be used in the diagnosis of asthma
 XX (increased levels) or IBD (reduced levels), also for monitoring
 XX treatment of these conditions. The ICACC proteins can be used:
 XX (1) to raise specific antibodies (Ab), useful: (a) as immunoassay
 XX reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
 XX (A) to disrupt bonding between ICACC and its ligands; and (iii) to
 XX identify modulators and binding partners. ICACC polynucleotides can be
 XX used to generate transgenic animals or recombinant cells, used to screen
 XX for antagonists, also as a source of therapeutic antisense agents or
 XX diagnostic probes (for quantifying mRNA expression, e.g. for
 XX identification of modulators).
 XX
 SO Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;

Query Match 100.0%; Score 248; DB 20; Length 2745;
 Best Local Similarity 100.0%; Pred. No. 4.8e-68;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ACCTGAAGCGGGAATTTCACGGGGGAGCTCATTAATCTGATGGACAGCTCTGGGG 60
 DB 2273 accggaagcggaatttcacgggggagctctcatlaactgagcttgagcagctccg 2332
 QY 61 ATGATTATGACCATGGAACAGCTCACAAGATATATCATTTGATTAAGTACGATTTCTTG 120
 DB 2333 atgattatgacctggaacagctcacaagatatactatcgtgaataagacagatattctg 2392
 QY 121 ATCTCAGAGACAAGTTCAATGAATCTTCAAGTGAATTAAGTCTGCTCTGATCCCAAGG 180
 DB 2393 atctcagagacaagttcaatgaatcttcaagtgaaatactactgctctcaccacaag 2452
 QY 181 AAGCCCACTCTGAGGAAGCTTTTGTTAACGAGAAACATTACTTTGAAAATGCGCA 240

Db 2453 aagcaactctgaggaagcttcttctttaaaccagaacattactttgaaatgca 2512
OY 241 CAGATCTT 248
|||||
Db 2513 cagatctt 2520

RESULT 5

AAH46124
ID AAH46124 standard; cDNA; 2825 BP.

AC AAH46124;

DT 11-SEP-2001 (first entry)

DE Human CLCA1 cDNA, SEQ ID NO:26.

XX Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
KM expression inhibition; antisense therapy; gene therapy;
KM chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
SS.

OS Homo sapiens.

XX Key Location/Qualifiers

FH 23..2767

FT CDS

FT /tag="a

FT /product="Human CLCA1"

FT /trans_except=" (pos:476..478, aa:lys)

XX WO200138530-A1.

PD 31-MAY-2001.

PF 22-NOV-2000; 2000WO-JP08232.

PR 24-NOV-1999; 99JP-0333479.

PR 27-APR-2000; 2000JP-0127589.

XX (TAKE) TAKEDA CHEM IND LTD.

XX PI Nakanishi A, Morita S;

DR WPI; 2001-355935/37.

DR P-PSDB; AAB73716.

PT New antisense nucleotide, useful for treatment and prevention of
PT bronchial asthma and chronic obstructive pulmonary disease -

PS Example 5; Page 92-94; 104pp; Japanese.

XX The invention relates to an antisense nucleotide targeted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the Gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
CC pulmonary disease. The present sequence represents human CLCA1 cDNA.
XX
SQ Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;

Query Match

Best Local Similarity 100.0%; Score 248; DB 22; Length 2825;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AACTGAAGCGGAATTCAGCGGCGAGTCATTATCTGACCTTGACAGCTCTGGGG 60
DB 2295 acctgaagcggaatcagcgggcgagtcattatctgactgagacagctcccg999 2354

OY 51 ATGATTATGACCATGAACAGCTCACAGTATATCATTCGATTAGTCAAGTATCTTG 120
|||||
Db 2335 atgattatgaccatgaaacagctcacaagtatatcatctgaataagtaagtatcttg 2414
OY 121 ATCTCAGACAGCAAGTTCATGATCTCTTCAAGTAATACTCTCTCATCCCAAG 180
|||||
Db 2415 atctcagacagtcacatgatactctcctcaagtgatactactgctcaccacaag 2474
OY 131 AAGCCAACTCAGAGAGCTTTTGTAAACCGAAGAACTTCTTTGAAATGGCA 240
|||||
Db 2475 aagccaaactcagaggaagcttcttctttaaaccagaacattactttgaaatgca 2534
OY 241 CAGATCTT 248
|||||
Db 2535 cagatctt 2542

RESULT 6

AAH34879
ID AAH34879 standard; cDNA; 2854 BP.

AC AAH34879;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:1961.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma; chromosome 1; ss.

OS Homo sapiens.

PD WO:00122920-A2.

PF 05-APR-2001.

PR 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (ETMA-) HUMAN GENOME SCI INC.

XX Fuwien SM, Barash SC, Birse CE, Rosen CA;

DR WFI; 2001-235357/24.

DR P-PSDB; AAG75474.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1; Page 3462-3463; 9803pp; English.

XX AA132943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG7789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query March	100.0%	Score 248;	DB 22	Length 2854;
Best Local Similarity	100.0%	Pred. No. 4.9e-68;		
Matches 248; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	ACGTGAAGCGGGAATTTCAACGGGGCAGTCTCATTAATCTGACTTGGACACTCTCTGGG	60
Db	2307	acctggaaggcggaattcaatccaggggcagctcctaatacttgacttgacaagctctctggg	2366
QY	61	ATGATTAATGACCATGGAGACAGCTCACAAATATATCATTCGAATTAAGTCAAGATATTCTTG	120
Db	2367	atgattatagcccatggaacagctccacaagtatatcatcttgaaatgaagtaacagtattcttg	2426
QY	121	ATCTCGAGAGCAAGTTCATCATGTAATCTCTCAAGTGAATCTACTGCTCTCATTCCTCAAGG	180
Db	2427	atctcgaagagcaagttccaatgaatcctctccaagtgaaatactactgctctcaccacaagg	2486
QY	181	AAGCCAACTCTGAGGAAGCTTTTTGTTTTAAACGAGAAACAATTACTTTGAAATGCA	240
Db	2487	aagccaactctggggaagctcttttgtttaacacagaaaacatacttlttgaaatgca	2546
QY	241	CAGATCTT 248	
Db	2547	cagatctt 2554	

RESULT 7

ID	AA	AF	81	787	standard;	CDNA;	2854	BP
1								

DT 12-JUN-2001 (first entry)

Human secreted protein gene 1 SEQ ID NO:11

KW Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW immunostimulant; cytostatic; cardiant; vascular; anti-angiotensic;
KW ophthalmological; neuroprotectant; nootropic; anticoagulant; vaccine;
KW antithesmer; antiparkinsonian; antimicrobial; vulnery; gene therapy
KW immune disorder; hyperproliferative disorder; cardiovascular disease;
KW cancer; angiotensin disorder; neurologic disorder; infectious disease;
KW wound healing; regeneration; chemotaxis; chromosome 1; ss

Homo sapiens.

PN WO200112775-A2.

PD 22-FEB-2001

PF 16-AUG-2000; 2000WO-US22325.
xy

PR 17-AUG-1999; 99US-0149182.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
YY

PI Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;
PI Birse CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;

DR WPI; 2001-147550/15.
DR P-PSDB; AAB74733.

PT Nucleic acids encoding 25 human secreted polypeptides, useful for preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease and diabetic retinopathy -

PS Claim 1; Page 441; 485pp; English.

CC AAF8187 to AAF8187 encode the human secreted proteins given in AAB747333
CC to AAB74772. Human secreted proteins can have activities based on the

CC tissues and cells they are expressed in. Example of activities include:
 CC immunomodulatory; antitumor; antileukemic; dermatological; immunosuppressive;
 CC anti-inflammation; anti-HIV; immunostimulant; cytostatic; cardiant;
 CC vasculuar; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
 CC anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial; and
 CC virolytic. Human secreted proteins can be used in gene therapy and
 CC vaccine. Human secreted protein nucleotide sequences (NAM1) and proteins
 CC (PEP1) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. For example, NAM1
 CC and PEP1 may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patients genome
 CC that affect the activity of proteins by expressing inactive proteins or
 CC to supplement the patients own production of polypeptides. Disorders that
 CC may be prevented, diagnosed and/or treated include immune disorders,
 CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
 CC autoimmune disorders, neurological disorders, infectious diseases and/or
 CC for promoting wound healing, regeneration and/or chemotaxis. AAR81778 to
 CC AAR81786 and AAR81792 represent sequences used in the exemplification of
 CC the present invention.

Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query/Match	100.0%	Score 248;	DB 22;	Length 2854;
Best Local Similarity	100.0%	Pred. NO.	4.9e-68;	
Matches 248;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

1 ACCTGAAGCGGGAATTCACGGGGGACGTCATTATCTGACTTGGACAGCTCTGGGG 60

[illegible]

A vertical ruler with markings from 0 to 10 cm. The markings are in millimeters, with numbers every centimeter. The ruler is oriented vertically, with the 0 mark at the top and the 10 cm mark at the bottom.

1. The first group of respondents (n = 10) was composed of students who had completed the course and were currently employed in a related field. 2. The second group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 3. The third group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 4. The fourth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 5. The fifth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 6. The sixth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 7. The seventh group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 8. The eighth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 9. The ninth group (n = 10) was composed of students who had completed the course and were currently employed in a related field. 10. The tenth group (n = 10) was composed of students who had completed the course and were currently employed in a related field.

[illegible]

100

— — — — —

Qy 241 CAGATCTT 248

Lb 2547 cagatctt 2554

RESULT 8

ID AAH35019 standard; cDNA; 3109 BP.

AC PA1135019

03 SEP-2001 (first entry)

Huñan colon cancer antigen encoding cDNA SEQ ID NO: 2101.

KV human; colon cancer; colon cancer antigen; diagnosis; detection;

XX

XX

XXI.

100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PR 03 NOV-1999; 99US-0163280.

PA (HJMA-) HUMAN GENOME SCI INC.

100

XX 12-JUL-2001.
 XX 29-DEC-2000; 2000MO-US35596.
 XX
 XX
 PR 30-DEC-1999; 99US-0476296.
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.

XX (CORI-) CORIXA CORP.

PI Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;
 PI King GE, Wang T, Jiang Y;

DR WPI: 2001-441847/47.

PT Colon tumor associated proteins and nucleic acids useful for the
 PT prevention, diagnosis and treatment of colonic cancer -

PS Claim 2; Page 425-426; 472pp; English.

XX The present invention describes colon tumour associated proteins (I) and
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate colon tumour associated protein (TCAP)
 CC expression, such as colonic cancer. For example, (I) and (II) may be
 CC used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of TCAPs by expressing inactive proteins or to supplement the
 CC patients own production of them. Additionally, (II) may be used to
 CC produce the TCAP proteins, by inserting the nucleic acids into a host
 CC cell culturing the cell to express the protein. (II) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC polymerase chain reaction (PCR) and hybridisation assays to detect and
 CC quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. (I) may
 CC also be used as antigens in the production of antibodies against TCAPs
 CC and in assays to identify modulators of TCAP expression and activity.
 CC Anti-(I) antibodies and antagonists may also be used to down regulate
 CC TCAP expression and activity. The anti-(I) antibodies may also be used
 CC as diagnostic agents for detecting the presence of TCAPs in samples
 CC (e.g. by enzyme linked immunosorbant assay (ELISA)). AA128460 to AA129512
 CC and AA124494 to AA124523 represent nucleotide and amino acid sequences
 CC given in the exemplification of the present invention.

XX Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;

Query Match 100.0%; Score 248; DB 22; Length 3311;
 Best Local Similarity 100.0%; Pred. No. 5.1e-68;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 ACCTGAAGGCGGAATTCACGGGGGAGTCATTAACTGACTTGGACAGCTCTGGG 60
 Db accctgaagcggaatcaccgggggagctcattcaatcgcagctgacagcctccg999 2663
 XX 61 ATGATTATGACATGGAACAGCTCACAAGTATATCATTCGAAATAGTACAAGTATCTTG 120
 Db atgattatgacatggaacagctcacaagtatcatcattcgaataagaatgattcttg 2743
 XX 121 ATCTCAGAGCAAGTTCATGAATCTTCTCAAGGATACTACTGCTCTCATCCCAAGG 180
 Db atctcagagcaagttcatagaatctctcaagtgaatactactgctctcatcccaaag 2803
 XX 181 AAGCCAACTCTGAGAGAGTCTTTTGTAAACCAAGAAACATACTTTTGAANAATGCA 240
 Db aagccaaactctgagagagctcttttgttaaaccaagaaacatattcttgaaaaatg 2863

CY 241 CAGATCTT 248
 Db 2864 cagatctt 2871

RESULT 11

AAH33285 standard; cDNA; 2867 BP.

AC AAH33285;

DT G3-SEP-2001 (first entry)

De Human colon cancer antigen encoding cDNA SEQ ID NO:341.

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

KN colorectal carcinoma; chromosome 1; ss.

CS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

XX 28-SEP-2000; 2000MO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rugen SM, Barash SC, Birse CE, Rosen CA;

DR WPI: 2001-235357/24.

XX P-PSDB; AAG73854.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -

PS Claim 1; Page 2452-2453; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.

GC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;

Query Match 99.4%; Score 246.4; DB 22; Length 2867;
 Best Local Similarity 99.6%; Pred. No. 1.6e-67;
 Matches 247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 1 ACCTGAAGGCGGAATTCACGGGGGAGTCATTAACTGACTTGGACAGCTCTGGG 60
 Db accctgaagcggaatcaccgggggagctcattcaatcgcagctgacagcctccg999 2370
 XX 61 ATGATTATGACATGGAACAGCTCACAAGTATATCATTCGAAATAGTACAAGTATCTTG 120

XX WO200055351-A1.
 PN 21-SEP-2000.
 XX 08-MAR-2000; 2000WO-US05883.
 PD 12-MAR-1999; 99US-0124270.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI: 2000-587534/55.
 DR P-PSDB: AAB53310.
 XX
 XX Colon cancer associated gene sequences, referred to as colon cancer
 PT antigens, useful for the treatment, prevention, and diagnosis of colon
 PT disorders such as colon cancer -
 XX
 PS Claim 1: Page 523; 2104pp; English.
 XX
 CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
 CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
 CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
 CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and
 CC vulnary, nephrotropic, antineoplastic and antibacterial activities, and
 CC can be used in gene therapy. The colon cancer antigen polynucleotides,
 CC proteins and antibodies to the proteins are useful for the prevention,
 CC treatment and diagnosis of colon disorders, such as colon cancer. The
 CC polynucleotides may be used in diagnostics and research, such as for
 CC chromosome identification, and as hybridisation probes. The proteins
 CC may also be used to prevent diseases such as neural disorders, immune
 CC system disorders, muscular disorders, reproductive disorders, infectious
 CC gastrointestinal disorders, wounds, renal disorders, infectious
 CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
 CC AAB54007 represent sequences used in the exemplification of the present
 CC invention.
 CC
 XX Sequence 1802 BP; 590 A; 348 C; 359 G; 485 T; 20 other;
 SQ

Query Match 52.3%; Score 129.8; DB 21; Length 1802;
 Best Local Similarity 75.6%; Pred. No. 5, 1e-31;
 Matches 158; Conservative 2; Mismatches 49; Indels 0; Gaps 0;

QY 33 ATTATCTGACTTGAGAGCTCTGAGGATGATTATGACCATGGAACGCTCACAAGTAT 92
 Db 1170 attattcttactgacagcagcagagataatttgatgttgaaagttcaacgklat 1229

QY 93 ATCATTCGAATAGTACAGTATTTGATCTCAGAGCAAGTTCATGAAATCTTCAA 152
 Db 1230 atyaaagaataagtcagatctctgataagaagagtttgatgctcttcaa 1289

QY 153 GTGAATCTACTGCTCTCATCCAAAGGAGCAACTCTGAGAACTTTTGTAA 212
 Db 1290 gtaataactactgactctgaccaaagagcgaactccaagaagaacttgcatttaaa 1349

QY 213 CCAGAAACATTACTTTTGAATAATGCGAC 241
 Db 1350 ccagaataatctcagaagaataatgcaac 1378

Search completed: April 4, 2002, 03:18:21
 Job time: 60253 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:04:07 ; Search time 22700.8 Seconds

(without alignments)
180.227 Million cell updates/sec

Title: US-09-049-696-14

Perfect score: 248
Sequence: 1 ACCTGAGGCGGAATTCAC.....TTGAAAATGCACAGATCTT 248

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenBankl :
1: gb_ba :
2: gb_htg :
3: gb_in :
4: gb_om :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vl :
15: em_ba :
16: em_fun :
17: em_hum :
18: em_in :
19: em_om :
20: em_ov :
21: em_or :
22: em_pat :
23: em_ph :
24: em_pl :
25: em_ro :
26: em_sts :
27: em_sy :
28: em_un :
29: em_vl :
30: em_htgo_hum :
31: em_htgo_inv :
32: em_htgo_rod :
33: em_htg_hum :
34: em_htg_inv :
35: em_htg_rod :
36: em_htg_other :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	248	100.0	878	6	195746	195746 Sequence 8
2	248	100.0	2826	9	AF127036	AF127036 Homo sapi
3	248	100.0	3311	6	AX193489	AX193489 Sequence
4	248	100.0	3311	9	AF039400	AF039400 Homo sapi
5	246.4	99.4	2022	9	AK024970	AK024970 Homo sapi
6	184.6	74.4	3079	4	AF095584	AF095584 Sus scrofa
7	169	68.1	35278	4	AF039401	AF039401 Homo sapi
8	169	68.1	113764	9	HSJ51E10	HSJ51E10 Homo sapi
9	169	68.1	164891	2	ALJ358950	ALJ358950 Homo sapi
10	149.4	60.2	2937	10	AB017156	AB017156 Mus muscu
11	129	52.0	3204	9	AF127035	AF127035 Homo sapi
12	129	52.0	3221	9	AK000072	AK000072 Homo sapi
13	129	52.0	3265	6	AK092338	AK092338 Sequence
14	127.4	51.4	1895	9	AK000138	AK000138 Homo sapi
15	102.2	41.2	140718	2	ALJ356270	ALJ356270 Homo sapi
16	94.6	38.1	2984	4	BTU36445	BTU36445 Bos taurus
17	88	35.5	3471	10	AF052746	AF052746 Mus muscu
18	87.4	35.2	3317	4	AF001261	AF001261 Bos tauru
19	87	35.1	3022	10	AF047838	AF047838 Mus muscu
20	82	33.1	3415	9	AF043976	AF043976 Homo sapi
21	79	31.9	2765	10	AF115852	AF115852 Mus muscu
22	79	31.9	3137	10	BC008147	BC008147 Mus muscu
23	78.4	31.6	3058	10	AF108501	AF108501 Mus muscu
24	65.4	26.4	140718	2	ALJ356270	ALJ356270 Homo sapi
25	65.4	26.4	175591	2	AC068071	AC068071 Homo sapi
26	60	24.2	549	6	AX054715	AX054715 Sequence
27	60	24.2	3604	9	AB026833	AB026833 Homo sapi
28	58.4	23.5	2832	9	AF127980	AF127980 Homo sapi
29	58.4	23.5	2970	9	AF043977	AF043977 Homo sapi
30	58.4	23.5	4077	6	AX054697	AX054697 Sequence
31	58.4	23.5	77948	2	AC018982	AC018982 Homo sapi
32	43.4	17.5	153552	9	AC009046	AC009046 Homo sapi
33	43.4	17.5	168172	9	AC010528	AC010528 Homo sapi
34	43.4	17.5	253694	2	AC084005	AC084005 Homo sapi
35	38.6	15.6	2652	9	AF104313	AF104313 Homo sapi
36	38.6	15.6	3086	9	AK001362	AK001362 Homo sapi
37	38.6	15.6	61344	2	AC073443	AC073443 Homo sapi
38	38.6	15.6	181254	2	AC016962	AC016962 Homo sapi
39	38.6	15.6	192247	2	AC013497	AC013497 Homo sapi
40	36.8	14.8	40738	8	YSCU9470	YSCU9470 Homo sapi
41	36.8	14.8	156424	2	AC025011	AC025011 Homo sapi
42	36.6	14.8	110	10	AF077303	AF077303 Rattus no
43	36.6	14.8	35583	3	CEC29P7	CEC29P7 Caenorhabd
44	36.4	14.7	14320	3	U00025	U00025 Caenorhabd
45	36.4	14.7	38725	3	M98552	M98552 Caenorhabd

ALIGNMENTS

RESULT 1
195746
LOCUS 195746 878 bp DNA
DEFINITION Sequence 8 from patent US 5733748.
ACCESSION 195746
VERSION 195746.1 GI:3940216
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 878)
AUTHORS Yu, G. and Rosen, C.
TITLE Colon specific genes and proteins
JOURNAL Patent: US 5733748-A 8 31-MAR-1998;
FEATURES
source 1. 878
location/Qualifiers
BASE COUNT 257 a 179 c 188 g 241 t 13 others
ORIGIN

PAT 01-DEC-1998

Db	2684	ATGATTATGACCATGGAACAGCTGCACAGTATATCATTCGATATAGTACAAAGTATTCCTGG	2743
OY	121	ATCTAGAGACAAGTTCAATGATATCTTTCAAGTAAATACTACTCTCATGCCAAGG	180
Db	2744	ATCTCAGAGACAAGTTCAATGATATCTTTCAAGTAAATACTACTCTCATGCCAAGG	2803
OY	181	AAGCCAACTCGAGGAAGTCTTTTGGTTAAACCGAANAACCTTACTCTTTGAAATGGCA	240
Db	2804	AAGCCAACTCGAGGAAGTCTTTTGGTTAAACCGAANAACCTTACTCTTTGAAATGGCA	2863
OY	241	CAGATCTT 248 	
Db	2864	CAGATCTT 2871	
RESULT	4		
AF039400			
LOCUS	AF039400	3311 bp	mRNA
DEFINITION	Homo sapiens calcium-dependent chloride channel-1 (hCLCA1)	14-DEC-1998	mRNA,
ACCESSION	AF039400		complete cds.
VERSION	AF039400.1	GI:4009457	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 3311) Gruber,A.D., Elble,R.C., Ji,H.L., Schreier,K.D., Fuller,C.M. and Pauli,B.U.		
TITLE	Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of Ca2+-activated Cl- channel proteins		
JOURNAL	Genomics 54 (2), 200-214 (1998)		
MEDLINE	99047526		
REFERENCE	2 (bases 1 to 3311) Gruber,A.D., Elble,R. and Pauli,B.U.		
AUTHORS	Direct Submission		
TITLE	Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA		
JOURNAL			
FEATURES			
SOURCE	Location/Qualifiers		
	1..3311		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="1p22-p31"		
	/tissue-type="small intestine"		
	1..3311		
	/gene="hCLCA1"		
	/note="transmembrane glycoprotein"		
	/codon_start=1		
	/product="calcium-dependent chloride channel-1"		
	/protein_id="AAC95428.1"		
	/db_xref="GI:4009458"		
	/translation="MGPFKSSVFLLIHLLEGALSNLIQLNNNGEGIVVAIDPNVPEDELITQIKDMTOASLYLEATGKRRFYFNVAIIPEMTKTADEVRRPELETKNAADIVAIAETPPGNDENRDEPTOMGCEKGRHLPPDPLAGKLAIFYGPGKAFVEMEARLRNVDFEYNNDEKFFYLSNGTQAVRCSAGITGTNNVYKKGCGGCTYKRTFNKVTGTYEKCEFLVLOSROTTERKASIMFAQHVDSIVERTQUNHNKEMPNQONCUNRSTWEVYLRSDSEKFTPTPTPTPPTPSLQIQIRIVCLVLDKSGSVATONRNLNRNLNOAGQLPLLOVLELSWGMGTFDSAAHVOSELIIQINSGBDITLAKRLPAAASGTSISGLRSRNLFTVIRKRYPTGSGEITVLLTDEEDNTISGCFNEVYOSGAIIHTVALGPSAOLEBELSKMTGGLQNYASPOVONNGLIIDAFGLASSGNAGVRSQIQLSEKGLTNSOMNGTIVYDSYTGKDTPLITTTTPTTPPOLIMDPSQKGGVYVDKNTFMATLIQIPGLTAVSTWTKYSLQSSQGTLLTIVTSRASNAFLPITVTSKTNKDTSKRPSPLVVYANTROGASPLRLRAAIVTALISVNGKVTLLLELDAGADATKDGVSRYFTYDINGRYSVKRALGQVNAARARVLPQOOSGALYIPGMILENDEIQNNPPELTKDDVQHKQYCFSTSGSGSFVASAARVAPLDLEPGQITDLKAEIHGGSILNTTVAAPGDYDHGAHNYIIRISITSLDLKQENASQVLENNFTALIPKKAENSEEVLFRKENTTPENGDTLFAIAYNVKDLKSRFLSNARVSLFIPQTPPTPEPSPDELSACPNNHINSTIPGHIILIKMKWIGELDLSTIA		

misc-feature /note="Cloning vector pME18SFL3"
 1..2022 /note="highly similar to AF127036 Homo sapiens
 calcium-activated chloride channel protein 1 (CaCC1) mRNA"
 BASE COUNT 612 a 472 c 453 g 485 t
 ORIGIN

Query Match 99.4%; Score 246.4; DB 9; Length 2022;
 Best Local Similarity 99.6%; Pred. No. 1.3e-58;
 Matches 247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ACCGTAGGCGGAAATTCACGGGGGAGTCTCATTAATGACTGGACAGCTCTGGGG 60
 DB 1486 ACCGTAGGCGGAAATTCACGGGGGAGTCTCATTAATGACTGGACAGCTCTGGGG 1545
 OY 61 ATGATTATGACATGAGACAGTCAAGTATATCATTTGAGTAAGTAAAGTATTTCTTG 120
 DB 1546 ATGATTATGACATGAGACAGTCAAGTATATCATTTGAGTAAGTAAAGTATTTCTTG 1605
 OY 121 ATCTGAGACAGATTCATGATCTCTCAAGTATCTACTGCTCTCATCCCAAGG 180
 DB 1606 ATCTGAGACAGATTCATGATCTCTCAAGTATCTACTGCTCTCATCCCAAGG 1665
 OY 181 AAGCCAACTCTGAGGAGTCTTTTGTAAACAGAAAACATTTACTTTGAAATGCGA 240
 DB 1666 AAGCCAACTCTGAGGAGTCTTTTGTAAACAGAAAACATTTACTTTGAAATGCGA 1725
 OY 241 CAGATCTT 248
 DB 1726 CAGATCTT 1733

RESULT 6

LOCUS AF095584 3079 bp mRNA MAM 13-OCT-2000
 DEFINITION Sus scrofa epithelial chloride channel protein (ABCC) mRNA,
 complete cds.
 ACCESSION AF095584
 VERSION AF095584.1 GI:6002645
 KEYWORDS
 SOURCE Pig.
 ORGANISM Sus scrofa

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 1 (bases 1 to 3079)
 Gaspar, K.J., Racette, K.J., Gordon, J.R., Loewen, M.E. and

TITLE
 Cloning a chloride conductance mediator from the apical membrane of
 porcine ileal enterocytes
 JOURNAL Physiol. Genomics (Online) 3 (2), 101-111 (2000)
 MEDLINE 20473747
 PUBMED 11015605

REFERENCE
 AUTHORS 2 (bases 1 to 3079)
 Gaspar, K.J., Gabriel, S.E., Racette, K.J. and Forsyth, G.W.
 TITLE Direct Submission
 JOURNAL Submitted (28-SEP-1998) Veterinary Physiological Sciences,
 University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4,
 Canada

FEATURES

SOURCE
 1..3079 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /tissue-type="ileal mucosa"
 1..3079
 /gene="ABCC"
 /gene="ABCC"
 133..2886
 /function="involved in exocrine chloride conductance"
 /codon_start=1
 /product="epithelial chloride channel protein"
 /protein_id="AAF00077.1"
 /db_xref="GI:6002646"

/translation="MGSEFRSSILFLVHLLEGASNSLIQNGNGEIVAIIDPNP
 EDERLIONIKNMVKASPYLFEEATEKREYFENKVALIPASMKAREYKPLEETKNA
 DVVTEPEPNDGQPYEPMONGCKGKIKYETDPVAGKRVILQYGPGRFVEHMAH
 LRKQVFNENYNNDEKFFYLNKKRQPYTCALIRGNIVLPQCGGSCYTRPCADRYTGL
 FQKREYFPDPOQSKASIMFAQSLDIVERCEKDNKKEAPNDQNCINRSTWEVL
 QDSDEKFTPTPTOPPAFTSLQIGRIVCLVKNKKEAPNDQNCINRSTWEVL
 LQTVQGMVGMVAFDSAAVYSELVQINSAEERDALALSLPTASGGTSICSGIRSA
 FTVIKKRPDSEIYLTLDGEDNTISACPEVKONGAIHTVALGSPAKRELELSQ
 MTGGLQYASQOAEENGLIDAFGLSSGNRAOSORIOLESQGLIONEMKNGTVVY
 DSTYQKDTLFLITLERKFLSPIPGVSGSGSDSPFLGKRNKMAVROVPTARVGMV
 KYSLQASQITLITVYSSRRSSATLPPTVYTKMKNDKTFSPVAVYTKHQDTPL
 RAKVTALLESNGKRVILELDNAGADATNDGSIYSFTAYDNGYSVKVALGG
 VNTPRRRAPPLMSGAMYTRGVIENGELIPNRPDINDKQVCSRTASGSEY
 ASDVKSPIPDLPFPCKITDILKAGIOGNLINTLWTAAGDDYDHGRADRYTIRISTNI
 LDLRKFNDSVOVWNTDLIPKEANSEVFEKPGIPETNGTDLFAVOAVDKTNKS
 EISNIAQVSLFLPEAPPEPPEPAPSLPCEPQVNSTIRGHIHKIMKMWLEIQL
 SIA"

BASE COUNT 892 a 758 c 755 g 674 t
 ORIGIN

Query Match 74.4%; Score 184.6; DB 4; Length 3079;
 Best Local Similarity 84.2%; Pred. No. 2.4e-41;
 Matches 208; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY 1 ACCGTAGGCGGAAATTCACGGGGGAGTCTCATTAATGACTGGACAGCTCTGGGG 60
 DB 241 ACCGTAGGCGGAAATTCACGGGGGAGTCTCATTAATGACTGGACAGCTCTGGGG 2470
 OY 61 ATGATTATGACATGAGACAGTCAAGTATATCATTTGAGTAAGTAAAGTATTTCTTG 120
 DB 2471 ATGATTATGACATGAGACAGTCAAGTATATCATTTGAGTAAGTAAAGTATTTCTTG 2530
 OY 121 ATCTGAGACAGATTCATGATCTCTCAAGTATCTACTGCTCTCATCCCAAGG 180
 DB 2531 ATCTGAGACAGATTCATGATCTCTCAAGTATCTACTGCTCTCATCCCAAGG 2590
 OY 181 AAGCCAACTCTGAGGAGTCTTTTGTAAACAGAAAACATTTACTTTGAAATGCGA 240
 DB 2591 AAGCCAACTCTGAGGAGTCTTTTGTAAACAGAAAACATTTACTTTGAAATGCGA 2650
 OY 241 CAGATCTT 247
 DB 2651 CAGATCTT 2657

RESULT 7

LOCUS AF039401 35278 bp DNA PRI 14-DEC-1998
 DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene,
 complete cds.
 ACCESSION AF039401
 VERSION AF039401.1 GI:4009459
 KEYWORDS
 SOURCE human.

REFERENCE
 AUTHORS Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 35278)
 Gruber, A.D., Elble, R.C., Ji, H.L., Schreuer, K.D., Fuller, C.M. and

TITLE
 Genomic cloning, molecular characterization, and functional
 analysis of human CLCA1, the first human member of the family of
 Ca2+-activated Cl- channel proteins
 JOURNAL Genomics 54 (2), 200-214 (1998)

REFERENCE
 MEDLINE 990473526
 2 (bases 1 to 35278)
 Gruber, A.D., Elble, R. and Pauli, B.U.
 TITLE Direct Submission

JOURNAL Submitted (19-DEC-1997) Department of Pathology, College of
 Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
 FEATURES
 SOURCE Location/Qualifiers
 1..35278


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
1.1617
/gene="hCLCA1"
1.33522
/gene="hCLCA1"
767.1265
/rpt_family="L1 transposable element"
/rpt_type="dispersed"
1587.1596
/gene="hCLCA1"
/feature="putative"
1618.1708
/gene="hCLCA1"
/number=1
join(1618..1708,1962..2383,6667..6807,6979..7126,9696,9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33522)
/gene="hCLCA1"
/product="calcium-dependent chloride channel-1"
1962..2383
/gene="hCLCA1"
/number=2
join(2222..2383,6667..6807,6979..7126,9696,9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766,31840..32079,32919..33310)
/gene="hCLCA1"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/db_xref="GI:4009460"
/translacion="MGEKSVFPIILHLLEGALSLSIQLNNGYEGVVAIDNPV
EDTLIQIKDWTQASLYLFETGKRFKFNVAALLIPETWKADYVRKLETVNA
DLVAVTEPDEPTQMGNGEGERIHLPDFIAKLALEYPOGAEFHEHNA
LHMVDEVDNDEKFTLSNGRIQAVRCSAGITGTNVKKGCGSCYTKRGTFRVGL
YKSGCFEVLQSRQTEKASIMPAOHDSIVIEFCGEONHNKEAPKOKCNLRSTMEVI
RSEDEPKTTPMTOPNPFTPSLQIGQIVICVLDKSSMATGNLNRNNGQLEL
LQTVELGSVWGVTFDPAHVOSELQINSQSDRDLARLPAAAGGTSICGLNSA
FTVIRKRYDQSEIYLIDGENTISGFENEKQGAIRHVALPSAOLELSEK
MTGQIYASDVONNGNLIDAGALSNGAVSORSIOLESKLTILQNSGMNVTYV
DSTVGKDTLFTLMTTOPPOILLMDPSQKGGFVVDKMTKMYLQIDPGAKGTWKY
SLOASSQTLITVTSRASNATLPIITVTSKTNKDSKPSPLVVVNIQOGASPIIRA
SVTALIESNGKVTLELDNGAGADATDDCYSSYFTTYDNGRYSKVRALGCVN
AARRYIPQSGALYIEGWIENDEIOMNPPRPITKDDVDHKOVCFSRISGGSFVAS
DVPNAIPDLPPGQITDINAEIHGSLINLTWTAGDDYDHGTAKYLIIRISTSLID
LADKFNESLOVNTTALIPKANSEEVFLKPEKITENGTDLFIAIOAVDKLSEI
SNIAVSLEIPQTPETSPDSTAPCPNIHINSTIPGIIHLIMMNGIELQSLIA"
6667..6807
/gene="hCLCA1"
/number=3
6979..7126
/gene="hCLCA1"
/number=4
9696..9801
/gene="hCLCA1"
/number=5
15454..15631
/gene="hCLCA1"
/number=6
18591..18809
/gene="hCLCA1"
/number=7
19783..20010
/gene="hCLCA1"
/number=8
22253..22427
/gene="hCLCA1"
/number=9

```

```

24522..24628
/gene="hCLCA1"
/number=10
26644..26859
/gene="hCLCA1"
/number=11
27447..27708
/gene="hCLCA1"
/number=12
28766..28936
/gene="hCLCA1"
/number=13
31840..32079
/gene="hCLCA1"
/number=14
32919..33522
/gene="hCLCA1"
/number=15
BASE COUNT 11245 a 6870 c 7230 g 9933 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
23
```


on the WORMPP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpp This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/RCR/Chr1>
 RP4-651E10 is from the library RCR1-4 constructed at the Roswell
 Park Cancer Institute by the group of Pieter de Jong. For further
 details see <http://bacpac.med.buffalo.edu/>
 VECTOR: pcypac2

This sequence is the entire insert of clone RP4-651E10.

FEATURES

Source

Location/Qualifiers
 1..113764
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="p22.3-31.1"
 /clone="RP4-651E10"
 /clone_lib="RCR1-4"
 BASE COUNT 34890 a 21989 c 22351 g 34534 t
 ORIGIN

Query Match 68.1%; Score 169; DB 9; Length 113764;
 Best Local Similarity 100.0%; Pred. No. 7.3e-37;
 Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 AGCTCAGCAAGTATATCATTCGTAATAGTACAGTATCTTGTGATCTAGAGCAAGTTCAA 139

Db 32622 AGCTCAGCAAGTATATCATTCGTAATAGTACAGTATCTTGTGATCTAGAGCAAGTTCAA 32681

QY 140 TGAATCTCTCAAGTGAATCTACTGCTCATCTCCCAAGAGCAACTCTGAGGAAGT 199

Db 32682 TGAATCTCTCAAGTGAATCTACTGCTCATCTCCCAAGAGCAACTCTGAGGAAGT 32741

QY 200 CTTTTCCTTTAAACAGAAAACATTTCTTTGAAATGCGACAGATCTT 248

Db 32742 CTTTTCCTTTAAACAGAAAACATTTCTTTGAAATGCGACAGATCTT 32790

RESULT 9

AL358950 164891 bp DNA HTG 23-JAN-2001
 LOCUS Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN
 DEFINITION PROGRESS ***, 10 unordered pieces.

ACCESSION AL358950.4 GI:12539689
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 164891)

Plumb.B.
 Direct Submission
 Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jan 26, 2001 this sequence version replaced gi:9988471.

COMMENT

Center: Sanger Centre
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk
 Project Information
 Center project name: ba444C12
 Summary Statistics
 Assembly program: XGAP; version 4.5
 Sequencing vector: plasmid; 108752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 161347 bases at least Q40
 Consensus quality: 162610 bases at least Q30
 Consensus quality: 163328 bases at least Q20

Insert size: 163991; sum-of-contigs
 Insert size: 163373; 6.6% error; agarose-fp
 Quality coverage: 5.19x in Q20 bases; sum-of-contigs Quality
 Coverage: 5.21x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently
 consists of 10 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 34236: contig of 34236 bp in length
 34237 34336: gap of 100 bp
 34337 47372: contig of 13036 bp in length
 47373 47472: gap of 100 bp
 47473 67025: contig of 19553 bp in length
 67026 67125: gap of 100 bp
 67126 71239: contig of 4114 bp in length
 71240 71339: gap of 100 bp
 71340 83903: contig of 12564 bp in length
 83904 84003: gap of 100 bp
 84004 91357: contig of 7354 bp in length
 91358 91457: gap of 100 bp
 91458 96452: contig of 4995 bp in length
 96453 96552: gap of 100 bp
 96553 148189: contig of 51637 bp in length
 148190 148289: gap of 100 bp
 148290 159778: contig of 11489 bp in length
 159779 159878: gap of 100 bp
 159879 164891: contig of 5013 bp in length.

FEATURES

Source

Location/Qualifiers
 1..164891
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-444C12"
 /clone_lib="RCR1-11.2"
 1..34236

misc-feature

/note="assembly-fragment:00027
 fragment_chain:1
 clone_end:r7
 vector_side:left"

misc-feature

/note="assembly-fragment:00312
 fragment_chain:1"

misc-feature

/note="assembly-fragment:01923
 fragment_chain:1"

misc-feature

/note="assembly-fragment:02094
 fragment_chain:1"

misc-feature

/note="assembly-fragment:02213
 fragment_chain:1"

misc-feature

/note="assembly-fragment:01292
 fragment_chain:1"

misc-feature

/note="assembly-fragment:01391
 fragment_chain:1"

misc-feature

/note="assembly-fragment:02045
 fragment_chain:1"

misc-feature

/note="assembly-fragment:00427
 fragment_chain:1"

misc-feature

/note="assembly-fragment:00773
 fragment_chain:1
 clone_end:sp6
 vector_side:right"

BASE COUNT 48381 a 32179 c 32357 g 51067 t 907 others
ORIGIN

Query Match 68.1%; Score 169; DB 2; Length 164891;

Best Local Similarity 100.0%; Pred. No. 7.5e-37;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 AGCTACAAATATATATCATTCGTAATAGTACAGATATCTTGATCTGAGACAGATTCAA 139
DB 49109 AGCTACAAATATATATCATTCGTAATAGTACAGATATCTTGATCTGAGACAGATTCAA 49050
QY 140 TGAATCTCTTCAAGTGAATACCTGCTCTCATCCCAAGAACCACTCTGAGAAGT 199
DB 49049 TGAATCTCTTCAAGTGAATACCTGCTCTCATCCCAAGAACCACTCTGAGAAGT 48990
QY 200 CTTTGTGTTAAACGAAACATTTACTTTGAAAAATGGACACAGATCTT 248
DB 48989 CTTTGTGTTAAACGAAACATTTACTTTGAAAAATGGACACAGATCTT 48941

RESULT 10

AB017156
LOCUS AB017156 2937 bp mRNA ROD 10-NOV-1999
DEFINITION Mus musculus gob-5 mRNA, complete cds.
ACCESSION AB017156
VERSION AB017156.1 GI:3721911
KEYWORDS GOB-5.
SOURCE Mus musculus adult intestine goblet cell cDNA to mRNA.
ORGANISM Mus musculus

REFERENCE 1 (sites)
AUTHORS Komiya, T., Tanigawa, Y. and Hirohashi, S.
TITLE Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice
JOURNAL Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)
MEDLINE 99160866
REFERENCE 2 (bases 1 to 2937)
AUTHORS Komiya, T., Tanigawa, Y. and Hirohashi, S.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1998) to the DDBJ/EMBL/GenBank databases. Tohru Komiya, ERATO, JST, Genome Asymmetry Group, Doi Biosymmetry Project, Tokodai 5-9-9, Tsukuba, Ibaragi 300-2635, Japan (E-mail: tkom@bioa.erauto.irc-net.co.jp, Tel:81-298-48-1515, Fax:81-298-47-8901)

FEATURES

source

1.2937
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="goblet cell"
/dev_stage="adult"
/tissue_type="intestine"
15..2756
/gene="gob-5"
15..2756
/gene="gob-5"
/codon_start=-1
/protein_id="BA033743.1"
/db_xref="GI:3721912"

CDS

gene

/translation="MESLKSYPVFLILHLEGLVLSLIQLNNNGYEGIVIALIDHNP
EDBALIHIKIDWVTOASPYLFEATGKRFYFNKVAIIIPESWKAPEYTRKLETFENNA
DVLVSTSPIDNDEPTEHIGAGGEGIRIHLPPDLAKGLTOYPODTEFHEHNA
FRMGVNEYNDKEFKYLSKGRQAVRCSAITKNGVRGCGGCTTGNGCYDVRVY
LYKNCVFVPPHONKASIMENONINSVEPCTEKNHNEAPNOCNLSASTWV
IOESDEKOTTPRTAPAPATPSLIDIGORIVLYLDKSGSMLNDRLRNMOASRIF
LIQVBOGSMVAVTFDSAAVYVSELKQLNSGADRLIKHLPTVSAGGTSISGLRT
APTVIKRKYPDGSEIVLTLDEDNFISCFDLYVKGSAIIRHVALGPAAKLEBLS
KMTGLOIYSSDOVONGVDAFAALSSGNAIAHOSIOLESRGVNLQNNOMNGSVI
VDSGKDLFLITTTTHPTPTIFIDPSGEONGFLIDTTTKVAYLOVPGTAVGFMK
YSIOASORLTLTIVTSRASATLPIITVIVYKNGKFKPSPTVYASIROGASPIIR
ASVTALIESVNGKTYVLELIDNGAGADKADKNGCYSTRFTAPANKRSYKIKALGCV
TSDORAAAPKRRNMTIDGWIIDEGEVRRMPPREPTIYVODKOLCFSRYSGGSFVATN

BASE COUNT 860 a 718 c 693 g 666 t

ORIGIN

Query Match 60.2%; Score 149.4; DB 10; Length 2937;

Best Local Similarity 75.3%; Pred. No. 1.7e-31;
Matches 186; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 1 ACCTGAAGCGGGAATTCACGGGGGCACTCATTTAACTGACTTGGACAGCTCTGGGG 60
DB 2287 ACCTGAAGCGGGAATTCACGGGGGCACTCATTTAACTGACTTGGACAGCTCTGGGG 2346
QY 61 ATGATTTATGACCCATGGAAGAGCTCAGATATATCATTTGAAATAGTAAAGTATCTTG 120
DB 2337 ATGATTTATGACCCATGGAAGAGCTCAGATATATCATTTGAAATAGTAAAGTATCTTG 2406
QY 181 ATCTCAGAGCAAGTGAATGATCTCTTCAAGTGAATGATCTCTCATCCCAAGG 180
DB 2407 ATCTCAGAGCAAGTGAATGATCTCTTCAAGTGAATGATCTCTCATCCCAAGG 2466
QY 181 AAGCCAACTCTGAGGAAGCTTTTGTGTTAAACCAAGAAACATTACTTTGAAATGCGCA 240
DB 2407 AAGCCAACTCTGAGGAAGCTTTTGTGTTAAACCAAGAAACATTACTTTGAAATGCGCA 2526
QY 241 CAGATCT 247
DB 2527 CAGATCT 2533

RESULT 11

AF127035
LOCUS AF127035 3204 bp mRNA PRI 11-AUG-1999
DEFINITION Homo sapiens calcium-activated chloride channel protein 2 (CaCC2)
ACCESSION AF127035
VERSION AF127035.1 GI:5726288
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3204)
AUTHORS Agnel, M., Vermal, T. and Culousoeu, J. M.
TITLE Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea
JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
REFERENCE 2 (bases 1 to 3204)
AUTHORS Agnel, M. and Culousoeu, J. M.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrières, Rueil-Malmaison 92500, France

FEATURES

source

1.3204
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="colon"
1..3204
/gene="CaCC2"
29..2782
/gene="CaCC2"
/note="bovine epithelial chloride channel homolog"
/codon_start=-1
/product="calcium-activated chloride channel protein 2"
/protein_id="AA048398.1"
/db_xref="GI:5726288"

CDS

gene

/translation="MGLFRGVFLVLCILHOSNTSPKLNNGFEDIVIVDPSVE
DEKTIQIDMTASTYLFEEATEKRFKNSILIPENKKNPOYKRRKHENKHAAD

VIVAPPLIPGDEPYTKQTFECGEYIHTPTDILLGKKNEXGPPKXLFVHEMHL
RMGVFDEINEDQPYFRANKSKIEKTRCSAGISGNRRYKCGGSCISRCHIDSTTL
YKDCQEPFPPVTPKASIMFQSIDVVECNKERTINOEPSONKCNRESTWEPI
SNSDEKFTIPMVPVPPVPSLKIQRVICALDSSGSGKDRNRNMAKHLL
LOJVENGSWGVHFDSTAIIVNKLIQIKSSDEBNITMAGLPTLGSTISCSGKIA
FOVGEHLSOLDSEVILLTDGEDNTASCSDEVKOSGAIYHFTALGRADEAVIENS
KITGSHRYVDEDMONNGILIDAFGLATSGNDLSOKSLOLESKGLTNSNMMDTYA
IDSTVGKTFEFLITMNSPSPISLMDPSGTMEPTVATSKMAVLSIPCTAKYGTVA
YNLOKAPDEFLITVTSRANSVPPITVAKANKNDYNSPSPMITAYELLOGVYV
LGANVTAFIESQNGTEVLELLDNGAGADSKNGVYSRYTATENGKRSKAKRAG
GANTRKLRPLNKAAYIPGVVNGELEAPPREDEIDOTYLEDGSRASGAGAV
VSOVSEPLPOVPSOTTDATYHEDKILITWAPDNDGVKQVYIIRISASIL
DLRSEFDALOVNTDLSKREANSKESPEPENSSEBATHIFAIKSIDKSNLSTK
VSNKQVTLFIPQANPDIDPPTPTPPPKSHNSGVNISTVLVIGSVIYVIFILS
TTI"

BASE COUNT 1098 a 594 c 633 g 879 t
ORIGIN

Query Match 52.0%; Score 129; DB 9; Length 3204;
Best Local Similarity 76.1%; Pred. No. 8.5e-26;
Matches 159; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 33 ATTAATCTGACTTGACAGCTCTCTGGGAGTATATGACATGGAACAGCTCAAGTAT 92
DB 2336 ATTATCTTACATGACAGACAGAGATATTTGATGTTGAAAGTTCAAGCTTAT 2395
QY 93 ATCATTGGAATTAAGTACAGATTTCTGATCTCAGAGACAGTCAATGATCTTCAA 152
DB 2396 ATCATAGATAAGTACAGATTTCTGATGATGAGAGAGTTTATGATGCTCTTCAA 2455
QY 153 GTGATACACTGCTCTCATCCCAAGAGCCACTGTGAGAGAGCTTTTGTAA 212
DB 2456 GTAAATACACTGATCTCTGACCAAGAGAGCCACTGTGAGAGAGCTTTGATTTAA 2515
QY 213 CCAGAAAACATTACTTTGAAAATGGCAC 241
DB 2516 CCAGAAAATATCTCAGAGAAAATGCAAC 2544

RESULT 12
AK000072 3221 bp mRNA PRI 22-FEB-2000
LOCUS Homo sapiens cDNA FLJ20065 f1s, clone COL01613, highly similar to
DEFINITION ECLC BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.
ACCESSION AK000072
VERSION AK000072.1 GI:7019922
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_11b:COL clone:COL01613.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (sites)
Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Ohayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
Unpublished (2000)
2 (bases 1 to 3221)
Sugano, S., Suzuki, Y., Ota, T., Ohayashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission
Submitted (15-FEB-2000) to the DDBJ/EMBL/Genbank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo, Department
of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@iims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry for Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology

Agency)
FEATURES
SOURCE location/Qualifiers
1..3221
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01613"
/clone_1b="COL"
/tissue="colon"
/note="cloning vector pME18FL3"
1..3221
/note="highly similar to ECLC BOVIN EPITHELIAL CHLORIDE
CHANNEL PROTEIN"
BASE COUNT 1105 a 600 c 634 g 882 t
ORIGIN

Query Match 52.0%; Score 129; DB 9; Length 3204;
Best Local Similarity 76.1%; Pred. No. 8.5e-26;
Matches 159; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 33 ATTAATCTGACTTGACAGCTCTCTGGGAGTATATGACATGGAACAGCTCAAGTAT 92
DB 2336 ATTATCTTACATGACAGACAGAGATATTTGATGTTGAAAGTTCAAGCTTAT 2409
QY 93 ATCATTGGAATTAAGTACAGATTTCTGATCTCAGAGACAGTCAATGATCTTCAA 152
DB 2410 ATCATAGATAAGTACAGATTTCTGATGATGAGAGAGTTTATGATGCTCTTCAA 2469
QY 153 GTGATACACTGCTCTCATCCCAAGAGCCACTGTGAGAGAGCTTTTGTAA 212
DB 2470 GTAAATACACTGATCTCTGACCAAGAGAGCCACTGTGAGAGAGCTTTGATTTAA 2529
QY 213 CCAGAAAACATTACTTTGAAAATGGCAC 241
DB 2530 CCAGAAAATATCTCAGAGAAAATGCAAC 2558

RESULT 13
AX092338 3265 bp DNA PAT 21-MAR-2001
LOCUS Sequence 69 from Patent WO0116318.
DEFINITION AX092338
ACCESSION AX092338
VERSION AX092338.1 GI:13444483
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3265)
Eaton, D.L., Pilyavoff, E., Gerlitsen, M.E., Goddard, A.,
Godowski, P.J., Grimaldi, C.J., Gurney, A.L., Watanabe, C.K. and
Wood, W.I.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
Patent: WO 0116318-A 69 08-MAR-2001;
Genentech, Inc. (US)
1..3265
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1159 a 596 c 632 g 878 t
ORIGIN

QY 33 ATTAATCTGACTTGACAGCTCTCTGGGAGTATATGACATGGAACAGCTCAAGTAT 92
DB 2332 ATTATCTTACATGACAGACAGAGATATTTGATGTTGAAAGTTCAAGCTTAT 2391
QY 93 ATCATTGGAATTAAGTACAGATTTCTGATCTCAGAGACAGTCAATGATCTTCAA 152

Query Match 52.0%; Score 129; DB 6; Length 3265;
Best Local Similarity 76.1%; Pred. No. 8.5e-26;
Matches 159; Conservative 0; Mismatches 50; Indels 0; Gaps 0;


```

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2392 ATCATAGATAAGTCAAGATATTTGATCTAGAGACAGTTTGTATGATGCTCTTCAA 2451
Oy 153 GTGATATCTACTGCTCTCATCCCAAGAGAACCACTCTGAGAGAGCTTTGTTTAA 212
Db 2452 GTAAATACACATGATCTGTCACCAAGAGAGCCCAACTCCAGAGAAAGCTTTGATTTAA 2511
Oy 213 CCAGAAACATTAATCTTTGAAATGCGCAC 241
Db 2512 CCAGAAATATCTCAGAGAAATGCAAC 2540

RESULT 14
AK000138 1895 bp mRNA PRI 22-FEB-2000
DEFINITION Homo sapiens cDNA FLJ20131 fis, clone COL06357.
ACCESSION AK000138
VERSION AK000138.1 GI:7020030
KEYWORDS cDNA capping; fis (full insert sequence);
SOURCE Homo sapiens colon cDNA to mRNA, clone_11b:COL clone:COL06357.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEO human cDNA sequencing project
JOURNAL Unpublished (2000)
TITLE 2 (bases 1 to 1895)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
JOURNAL Direct Submission
COMMENT Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo, Department
of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ems.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing; Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source 1. 1895
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL06357"
/clone_11b="COL"
/tissue_type="colon"
/feature="cloning vector pME18SFL3"
66. 1475
/feature="unnamed protein product"
/codon_start=1
/protein_id="BAA90969.1"
/db_xref="GI:7020031"
/translation="MSKTTGSGHFVSDENONNGJIDARFALTSQNTLSOKSLOLES
KGLTNSNANMNDYVITDSYVCKDPFELLTNSLSPSISLMPBSGIMTNFYDASK
MAYLISPGTAKYGTAVYNLOAKANPELTITVTSRAANSVPPITVNAKNNDVNSFP
SPMIVAEILQGYEVLGANTVAFIESQNGHEVELLDNGAGADSFKNQGVSRVET
AYTENGREYSLKVRAGANTARLKLRPBNRAAYIFGVVNGEIEANPREDEDTQ
TLEDESRFASGAFVSVQSPSLPDPQVQITPDLDATVEDKILITWPTAGDVED
VGKQRYIIRISASILDRSFDALQVTTDLSPEANSKSFARPKPNISENATH
IFIAKSIDKSNLTSKVSNIQAVTLFIIPANDDIDPPTPTPKSNKSNVSNISTL
VLSVIGSVIVNFIISTT"
BASE COUNT 658 a 352 c 357 g 528 t
ORIGIN
Query Match 51.4%; Score 127.4; DB 9; Length 1895;
Best Local Similarity 75.6%; Pred. No. 2.3e-25;

```

```

Matches 158; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
Oy 153 ATTAATCTGACTGGACAGCTCTGGGATGATATATACCATGAGACAGCTCACAAATAT 92
Db 1069 ATTATCTTCATGACGACAGCAGCAGCATATTTTATTTATTTGGAAGATTCACGTTAT 1088
Oy 153 ATCATTCGATTAAGTACATATTTCTGATCTCAGACACAACTCAATGATCTCTTCAA 152
Db 1069 ATCATAGAAATTAAGTACATATTTCTGATCTCAGACACAACTCAATGATCTCTTCAA 1148
Oy 153 GTGATATCTACTGCTCTCATCCCAAGAGAACCACTCTGAGAGAGCTTTGTTTAA 212
Db 1069 GTAAATACACATGATCTGTCACCAAGAGAGCCCAACTCCAGAGAAAGCTTTGATTTAA 1208
Oy 153 CCAGAAACATTAATCTTTGAAATGCGCAC 241
Db 1209 CCAGAAATATCTCAGAGAAATGCAAC 1237

RESULT 15
AL356270/c
LOCUS AL356270
DEFINITION Homo sapiens chromosome 1 clone RP5-873015, *** SEQUENCING IN
PROGRESS ***
ACCESSION AL356270
VERSION AL356270.3 GI:9797470
KEYWORDS HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 140718)
AUTHORS Pavitt,R.
JOURNAL Direct Submission
COMMENT Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Aug 12, 2000 this sequence version replaced g1:9213800.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
Project information
Center project name: dj873015
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M7815; 12% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 129686 bases at least Q40
Consensus quality: 134097 bases at least Q20
Consensus quality: 136445 bases at least Q20
Insert size: 138618; sum-of-contigs
Insert size: 144145; 13.2% error; agarose-fp
Quality coverage: 3.10x in Q20 bases; sum-of-contigs Quality
coverage: 3.21x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 3292: contig of 3292 bp in length
* 3293 3392: gap of 100 bp
* 3393 10168: contig of 6776 bp in length
* 10169 10268: gap of 100 bp
* 10269 21700: contig of 11432 bp in length
* 21701 21800: gap of 100 bp
* 21801 29662: contig of 7862 bp in length

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:54:33 ; Search time 521.98 Seconds
(without alignments)
107.603 Million cell updates/sec

Title: US-09-049-696-14

Perfect score: 248
Sequence: 1 ACCTGAAGCGCGAAATTCAC.....TTGAAATGCGACAGATCTT 248

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 segs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	248	100.0	878	1	US-08-469-667-8
2	248	100.0	878	5	PCT-US95-07289-8
3	129	52.0	618	4	US-09-385-982-24
4	128.2	51.7	595	4	US-09-385-982-25
5	92.8	37.4	742	4	US-09-385-982-33
6	44.4	17.9	576	4	US-09-385-982-23
7	33	13.3	1984	1	US-07-885-970A-25
8	33	13.3	1985	1	US-08-298-687A-25
9	33	13.3	1885	1	US-08-298-829-25
10	30.8	12.4	864	4	US-08-953-073A-1
11	30	12.1	4362	2	US-08-455-073A-1
12	29	11.7	1554	2	US-08-031-538-8
13	29	11.7	1730	1	US-07-817-920-1
14	29	11.7	1730	1	US-08-117-006-1
15	29	11.7	1730	1	US-08-216-594-1
16	29	11.7	1730	5	PCT-US93-00149-1
17	29	11.7	11283	3	US-08-603-753D-3
18	29	11.7	11283	4	US-09-099-753-3
19	29	11.7	11283	4	US-08-986-106-3
20	28.6	11.5	2273	3	US-08-714-918-40
21	28.6	11.5	2273	4	US-09-265-315-40
22	28.6	11.5	2273	4	US-09-265-315-40
23	28.6	11.5	2273	4	US-09-266-417-40
24	28.4	11.5	2275	2	US-08-743-637B-2
25	28.4	11.5	2275	2	US-08-526-840B-2
26	28.4	11.5	3828	5	PCT-US93-10500-1
27	28.4	11.5	4190	2	US-08-488-706-3

28	28.4	11.5	7721	3	US-08-772-270A-14	Sequence 14, Appl
29	28.2	11.4	676	1	US-08-259-745A-41	Sequence 41, Appl
30	28.2	11.4	676	1	US-08-259-745A-42	Sequence 42, Appl
31	28.2	11.4	676	1	US-08-259-745A-43	Sequence 43, Appl
32	28.2	11.4	1761	3	US-09-033-055A-3	Sequence 3, Appl
33	28.2	11.4	2730	1	US-08-339-129-1	Sequence 1, Appl
34	28	11.3	378	4	US-09-328-111-191	Sequence 191, App
35	28	11.3	971	5	PCT-US91-06234A-5	Sequence 5, Appl
36	28	11.3	1659	1	US-08-231-729B-3	Sequence 3, Appl
37	28	11.3	1659	1	US-08-231-729B-4	Sequence 4, Appl
38	28	11.3	2063	1	US-08-231-729B-5	Sequence 5, Appl
39	28	11.3	2063	3	US-08-781-891-79	Sequence 79, Appl
40	28	11.3	87350	3	US-09-048-889-9	Sequence 9, Appl
41	27.8	11.2	1496	3	US-08-475-844-20	Sequence 20, Appl
42	27.6	11.1	520	2	US-08-475-844-20	Sequence 20, Appl
43	27.6	11.1	530	5	PCT-US95-08429-20	Sequence 20, Appl
44	27.4	11.0	4526	1	US-07-855-412B-4	Sequence 4, Appl
45	27.4	11.0	4526	2	US-08-308-887A-4	Sequence 4, Appl

ALIGNMENTS

RESULT: 1

US-08-469-667-8
Sequence 8, Application US/08469667

Patent No. 5733748

GENERAL INFORMATION:

APPLICANT: Yu, Guo-Liang

INVENTOR: Rosen, Craig

TITLE OF INVENTION: Colon Specific Genes and Proteins

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carella, Byrne, Bain, Gillfillan, Cecchi,

ADDRESS: Stewart & Olstein

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: NJ

COUNTRY: USA

ZIP: 07068-1739

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,667

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Ferraro, Gregory D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-435

TELEPHONE: 201-994-1744

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 878 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2..685

US-08-469-667-8

Query Match 100.0% Score 248; DB 1; Length 878;
Best Local Similarity 100.0%; Pred. No. 1.4e-68;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCTGAAGCGGAATTCACGGGGCACTCTCATTAATCTGACTTGACAGCTCTCTGGGG 60
 DB 306 ACCTGAAGCGCGGAATTCACGGGGCACTCTCATTAATCTGACTTGACAGCTCTCTGGGG 365
 QY 61 ATGATTATGACATGACAGCTCAGAGTATATCATTCGATTAAGTATCAAGTATCTTG 120
 DB 366 ATGATTATGACATGACAGCTCAGAGTATATCATTCGATTAAGTATCAAGTATCTTG 425
 QY 121 ATCTGAGACAGTTCATTAATCTCTGAGTATCTGCTCTCATCCCAAGG 180
 DB 426 ATCTGAGACAGTTCATTAATCTCTGAGTATCTGCTCTCATCCCAAGG 485
 QY 181 AAGCCACTCTGAGAGCTTTTGTGTTAAACCAAGAACTCTTTGAAATGGCA 240
 DB 486 AAGCCACTCTGAGAGCTTTTGTGTTAAACCAAGAACTCTTTGAAATGGCA 545
 QY 241 CAGATCTT 248
 DB 546 CAGATCTT 553

RESULT 2

PCT-US95-07289-8
 Sequence 8, Application PC/TUS9507289

GENERAL INFORMATION:
 APPLICANT: Yu, Guo-Liang
 APPLICANT: Rosen, Craig
 TITLE OF INVENTION: Colon Specific Genes and Proteins
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Carella, Byrne, Balin, Gillfillan, Cecchi,
 STREET: 6 Becker Farm Road
 CITY: Roseland
 STATE: NJ
 COUNTRY: USA
 ZIP: 07068-1739
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/07289
 FILING DATE: 06-JUN-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Ferraro, Gregory D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 325800-265
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 878 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 2..685
 PCT-US95-07289-8

Query Match 100.0%; Score 248; DB 5; Length 878;
 Best Local Similarity 100.0%; Pred. No. 1,4e-68;
 Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCTGAAGCGGAATTCACGGGGCACTCTCATTAATCTGACTTGACAGCTCTCTGGGG 60
 DB 306 ACCTGAAGCGCGGAATTCACGGGGCACTCTCATTAATCTGACTTGACAGCTCTCTGGGG 365

QY 61 ATGATTATGACATGACAGCTCAGAGTATATCATTCGATTAAGTATCAAGTATCTTG 120
 DB 366 ATGATTATGACATGACAGCTCAGAGTATATCATTCGATTAAGTATCAAGTATCTTG 425
 QY 121 ATCTGAGACAGTTCATTAATCTCTGAGTATCTGCTCTCATCCCAAGG 180
 DB 426 ATCTGAGACAGTTCATTAATCTCTGAGTATCTGCTCTCATCCCAAGG 485
 QY 181 AAGCCACTCTGAGAGCTTTTGTGTTAAACCAAGAACTCTTTGAAATGGCA 240
 DB 486 AAGCCACTCTGAGAGCTTTTGTGTTAAACCAAGAACTCTTTGAAATGGCA 545
 QY 241 CAGATCTT 248
 DB 546 CAGATCTT 553

RESULT 3

US-09-385-982-24/c
 Sequence 24, Application US/09385982

GENERAL INFORMATION:
 APPLICANT: ENDEGE, WILSON O., ET AL.
 TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 FILE REFERENCE: CCDNA-260XX
 CURRENT FILING DATE: 1999-08-30
 EARLIER APPLICATION NUMBER: 09/328,111
 EARLIER FILING DATE: 1999-06-08
 EARLIER APPLICATION NUMBER: 60/117,393
 EARLIER FILING DATE: 1999-01-27
 EARLIER APPLICATION NUMBER: 60/098,639
 EARLIER FILING DATE: 1998-08-31
 NUMBER OF SEQ ID NOS: 544
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 24
 LENGTH: 618
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc-feature
 LOCATION: (1)...(618)
 OTHER INFORMATION: n = A,T,C or G
 US-09-385-982-24

Query Match 52.0%; Score 129; DB 4; Length 618;
 Best Local Similarity 76.1%; Pred. No. 1.9e-31;
 Matches 159; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 33 ATTAATCTGACTTGACAGCTCTCTGGGATTAATGACATGAGAACGCTCACAAGTAT 92
 DB 441 ATTAATCTGACTTGACAGCTCTCTGGGATTAATGAGAACGCTCACAAGTAT 382
 QY 93 ATCATCTGATTAAGTATCTCTGAGTATCTGAGTATCTGAGTATCTGAGTATCTG 152
 DB 381 ATCATCTGATTAAGTATCTCTGAGTATCTGAGTATCTGAGTATCTGAGTATCTG 322
 QY 153 GTGAATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 212
 DB 321 GTGAATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 262
 QY 213 CCAGAAAACATTAATCTTTGAAAATGGCAC 241
 DB 261 CCAGAAAACATTAATCTTTGAAAATGGCAC 233

RESULT 4
 US-09-385-982-25
 Sequence 25, Application US/09385982
 Patent No. 6262334

GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 25
LENGTH: 595
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(595)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-25

Query Match 51.7%; Score 128.2; DB 4; Length 595;
Best Local Similarity 75.1%; Pred. No. 3,4e-31;
Matches 157; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 33 ATTAATCTGACTTGACAGCTCTCTGGGATGATTATGACCATGACAGCTCACAAGTAT 92
DB 210 attattcttaacatgagcagcagagataatttgatgtgaaagttcaacgntat 269
QY 93 ATCATTCGATTAAGTACAGTATTTCTTGATCTGAGACAGACAGTTCATGATCTTCAA 152
DB 270 atcataagaataagatgacatcttctgataagagagcttlnatgattgctctca 329
QY 153 GTGAATACACTGCTCTGCATCCCAAGAGAGCAAGCTCTGAGAGAGCTTTTGTAA 212
DB 330 gtaataactactgactgtcaccagaagagagccaactcccaagaagaagcttgcntttaa 389
QY 213 CCAGAAACATTTCTTTGAAAATGCGAC 241
DB 390 ccagaaataatctcagaagaagaatgcaac 418

RESULT 5
US-09-385-982-33
Sequence 33, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 742
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(742)

OTHER INFORMATION: n = A,T,C or G
US-09-385-982-33
Query Match 37.4%; Score 92.8; DB 4; Length 742;
Best Local Similarity 74.1%; Pred. No. 4.2e-20;
Matches 143; Conservative 0; Mismatches 48; Indels 2; Gaps 2;
QY 33 ATTAATCTGACTTGACAGCTCTCTGGGATGATTATGACCATGACAGCTCACAAGTAT 92
DB 210 attattcttaacatgagcagcagagataatttgatgtgaaagttcaacgntat 269
QY 93 ATCATTCGATTAAGTACAGTATTTCTTGATCTGAGACAGACAGTTCATGATCTTCAA 152
DB 270 atcataagaataagatgacatcttctgataagagagcttlnatgattgctctca 329
QY 153 GTGAATACACTGCTCTGCATCCCAAGAGAGCAAGCTCTGAGAGAGCTTTTGTAA 212
DB 310 gtaaat-ctactgactc-gacccaagagagccaactccaagaagaagcttgcattaac 387
QY 213 CCAGAAACATTTA 225
DB 318 caaaaaatattta 400

RESULT 6
US-09-385-982-23
Sequence 23, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(576)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-23

Query Match 17.9%; Score 44.4; DB 4; Length 576;
Best Local Similarity 74.5%; Pred. No. 5.1e-05;
Matches 82; Conservative 0; Mismatches 26; Indels 2; Gaps 2;
QY 33 ATTAATCTGACTTGACAGCTCTCTGGGATGATTATGACCATGACAGCTCACAAGTAT 92
DB 42 attattcttaacatgagcagcagagataatttgatgtgaaagttcaacgntat 501
QY 93 ATCATTCGATTAAGTACAGTATTTCTTGATCTGAGACAGACAGTTCATGATCTTCAA 142
DB 502 atcataagaataa-tgccagattcttga-ctaaagagacagtttgatga 549

RESULT 7
US-07-885-970A-25/c
Sequence 25, Application US/07885970A
Patent No. 5495070
GENERAL INFORMATION:


```

TITLE OF INVENTION: PLANTS FOR ALTERED FIBER
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nicholas J. Seay, Quarles & Brady
STREET: P.O. Box 2113, First Wisconsin Plaza
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53701

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/298,687A
FILING DATE:

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/617,239
FILING DATE: 21-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/253,243
FILING DATE: 04-OCT-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27,386
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 283-2478
TELEFAX: (608) 251-5139
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1985 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
TOPOLOGY: linear
HYBOTHEITCAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium barbadense
STRAIN: Sea Island
IMMEDIATE SOURCE:
LIBRARY: EMBL-SI
CLONE: SIH6
US-08-298-687A-25

Query Match 13.3%; Score 33; DB 1; Length 1985;
Best Local Similarity 53.5%; Pred. No. 0.3;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0.

QY 90 TATATCATTGCAATAGTACGATGTTTCTTGATCTCAGAGACAGTTCATTAATCTCTT 149
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1818 TATATGATTCGATTAATTAACCAACCAACATATATCCAAACAAATTAATTCCTT 1759
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 130 CAAGTATACACTACTGCTCTCATCCCAAGGAGCCAACTTGAGGAAGTCTTTTGGTT 209
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1758 CAATACCAAAAAAAGAGTCTCTCAAAAGGTAAAAAATTAATTAATCTCCAAAGAT 1699
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 210 AAACCGAA 218
   |||||
Db 1698 ACAATATATA 1690
   |||||

RESULT 9
US-08-298-829-25/c
Sequence 25, Application US/08298829
Patent No. 5620882
GENERAL INFORMATION:
Applicant: John, Maliyakal E.
TITLE OF INVENTION: GENETICALLY ENGINEERING COTTON
PLANTS FOR ALTERED FIBER

```



```

NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nicholas J. Seay, Charles & Brady
STREET: P.O. Box 2113, First Wisconsin Plaza
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53701

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/298,829
FILING DATE: 19-OCT-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/885,970
FILING DATE: 18-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/617,239
FILING DATE: 21-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/253,243
FILING DATE: 04-OCT-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27,386
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 283-2478
TELEFAX: (608) 251-5139
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1985 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium barbadense
STRAIN: Sea Island
IMMEDIATE SOURCE:
LIBRARY: EMBL-SI
CLONE: SIH6

```

```

Query Match          13.3%; Score 33; DB 1; Length 1985;
Best Local Similarity 53.5%; Pred. NO. 0.3;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0.

QY 90 TATATCATTCGAATTAAGTACAAAGTATTTCTTGATCTCAGAGACCAAGTTCATGATCTCTT 149
      ||||| | ||||| | ||| | | | | | | | | | | | | | | | | | | | | | | |
Db 1818 TATATGAATCGAAATTAATCAACCAACCAATATATTCACCAACAAATTAACCAATTAATTCCTT 1759

QY 150 CAAGGATTAAGTACTGCTCTCATCCCAAGCAAGCAACTCGAGGAAGTCTTTTGTGT 209
      ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||
Db 1758 CAATTAACCAAAAAAAGAGTGCTCTCAAAAAGGTAAATAAAATTAATTAATGTCTCAAAAGAT 1699
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 210 AAACCCAGAA 218
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1698 ACAATATATA 1690

RESULT 10
US-08-953-326-12
: Sequence 12, Application US/08953326
: Patent No. 6251872
: GENERAL INFORMATION:
: APPLICANT: Bardel, Anthony F.

```

APPLICANT: Ganta, Roman R.
APPLICANT: McGuire, Travis C.
APPLICANT: Butridge, Michael J.
APPLICANT: Nyika, Aceme
APPLICANT: Rutangirwa, Fred R.
APPLICANT: Mahan, Susan M.
TITLE OF INVENTION: Nucleic Acid Vaccines Against Rickettsial Diseases Of
TYPE OF INVENTION: Animals and Humans
FILE REFERENCE: UF-167C1
CURRENT APPLICATION NUMBER: US/08/953,326
CURRENT FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/953,326
EARLIER FILING DATE: 1997-10-17
EARLIER APPLICATION NUMBER: 08/733,230
EARLIER FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 864
TYPE: DNA
ORGANISM: Ehrlichia canis
US-08-953-326-12

Query/Match 12.4% Score 30.8; DB 4; Length 864;
Best Local Similarity 55.7% Pred. No. 1.1;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

56 TGGGATGATTATGACCATGGAAGCTTCACACAGTATATCAATGCATAATGACAAGTAT 115
pb | | | | | | | | | | | | | | | | | | | | | | | | | | | |
44 ttagcggaattggtacactgcatacaaacggataagtttgtaacctcggaaaatgaaggltt 523

QY 116 TCTTGATCTCACAGACAACTTCATCATGTAATCTTTCAAGTGAATACT 161
Do | | | | | | | | | | | | | | | | | | | | | | | | | | | |
54 acttgagctccatttatgtttaacgcatgtatgacaatacaact 569

RESULT 11
US-08-145-073A-1
Sequence 1, Application US/08455073A
Patent No. 5876949

GENERAL INFORMATION:
APPLICANT: Gideon Dreyfuss
APPLICANT: Mikiko C. Sioml
APPLICANT: Yan Zhang
TITLE OF INVENTION: Fragile X Related Proteins, Compositions And Methods
TITLE OF INVENTION: Of Making And Using The Same
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5876949ris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,073A
FILING DATE: 31-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: OPN-2201
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 4362 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 220..2118
US-08-455-073A-1

Query Match 12.1%; Score 30; DB 2; Length 4362;
Best Local Similarity 54.5%; Pred. No. 3.5;
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 132 AAGTCAATGATCTCTGATGAGTACTACTGCTGCTGATCCCAAGAGGACGACTCT 191
DB 3915 ATGTTGATGCAATCCCTTACAAATGATGCTTTAAATTTAAGTAGAGAAAGAAATCT 3974
QY 192 GAGGAAGTCTTTTGTGTTAAACCAGAAACATTTCTTTGAAATGGCAC 241
DB 3975 ATAGAAAGTCTCTGTGTACAAATATGTAATGTTACATTTGAAATTTTAC 4024

RESULT 12

US-08-031-538-8
Sequence 8, Application US/08031538
Patent No. 5968817

GENERAL INFORMATION:

APPLICANT: Sutcliffe, J Gregor
APPLICANT: Erlander, Mark G
APPLICANT: Lovenberg, Timothy W
TITLE OF INVENTION: HUMAN SEROTONIN RECEPTORS, DNA ENCODING
TITLE OF INVENTION: THE RECEPTORS, AND USES THEREOF
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESS: Patent Counsel
STREET: 10666 No. 5968817th Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/031,538
FILING DATE: 19930315
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSM5099P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 1554 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-031-538-8

Query Match 11.7%; Score 29; DB 2; Length 1554;
Best Local Similarity 52.8%; Pred. No. 4.8;

Matches 85; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 137 AAGTATATCATTCGATATAGTATGATCAATGATCTTGATCTGACAGACAGATTCATGATC- 145
DB 139 AAACATACATACACTTTTAAATATTTCTGAAAGAGAGAAAGTTCTTGAAGCCT 148
QY 146 TCTTCAAGTGAATACTACTCTCATCCCAAGAAAGCAACTGAGGAAGTCTTTT 205
DB 149 TCTGGAAGTCTTTTCTCTCTCCCTGTTACAGGATCATCTTTTACACTATATTATTC 208

QY 206 GTTTAACCAGAAACATTTCTTTGAAATGGCACAGATC 246

DB 209 TTTTAAACAAAGAAATGATTTCTTAAATTCATCTGATC 249

RESULT 13

US-07-817-920-1
Sequence 1, Application US/07817920
Patent No. 5360735

GENERAL INFORMATION:

APPLICANT: Weinshank, Richard L
APPLICANT: Branchek, Theresa
APPLICANT: Bartig, Paul R
TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/817,920
FILING DATE: 19920108
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1795/39318
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525
FILING DATE: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1730 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
INVENTIVE SOURCE:
LIBRARY: human lymphocyte genomic
CLONE: h116a
FEATURE:
NAME/KEY: CDS
LOCATION: 616..1713
NAME/KEY: mat_peptide
LOCATION: 616..1713
US-07-817-920-1

Query Match 11.7%; Score 29; DB 1; Length 1730;

Thu Apr 4 09:27:30 2002

us-09-049-696-14.rn

Page 8

Query Match	11.7%	Score 29;	DB 1;	Length 1750;
Best Local Similarity	52.8%	Pred. No. 5;		
Matches	85;	Mismatches	75;	Indels 1;
				Gaps 1.

[illegible]

Search completed: April 3, 2002, 20:54:36
Job time: 41834 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:38 ; Search time 16681 Seconds
(Without alignments)
146.231 Million cell updates/sec

Title: US-09-049-696-13

Perfect score: 227

Sequence: 1 GTGGGGGCTGTGAGAGAGT.....AATGCTCCACTGACTGACT 227

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	227	100.0	455	11	BF083321 RCL-CT024
2	227	100.0	598	11	BE927204 RCL-CT024
3	225.4	99.3	528	11	BF083270 RCL-CT024
4	168.4	74.2	315	11	BF083396 RCL-CT043
5	107.4	47.3	734	10	AI660957 WT20C08.X
6	104.8	46.2	2915	12	AK007466 Mus muscu
7	103.2	45.5	469	10	AA611564 VO85D05.T
8	101.6	44.8	2933	12	AK008659 Mus muscu
9	97.8	43.1	653	10	AW361534 OV2-CT026
10	94.8	41.8	794	11	BM320001 RST39773
11	94.8	41.8	813	11	BG216984 RST36685
12	90.4	39.8	528	11	BG965506 602830872

13	89.2	39.3	449	10	AA759816
14	85.2	37.5	437	10	BE136979
15	85.2	37.5	527	10	BE137034
16	84.2	37.1	911	11	BF578693
17	79.4	35.0	974	11	BG962464
18	78.2	34.4	420	10	AI641868
19	71	31.3	450	10	AA734088
20	62.8	27.7	646	10	AA691335
21	61.2	27.0	717	11	BE850924
22	61.2	27.0	824	11	BI109423
23	61.2	27.0	844	11	BG174331
24	59.4	26.2	793	10	BE913856
25	58.2	25.6	739	10	AA541829
26	58	25.6	774	11	BF583551
27	56.6	24.9	653	10	AM212836
28	55.8	24.6	524	10	AA871630
29	54.8	24.1	367	11	BF523821
30	54.8	24.1	466	10	AA170458
31	54.8	24.1	468	10	AI640041
32	53.8	23.7	770	10	BE738113
33	52	22.9	812	11	BG968878
34	48.2	21.2	633	10	AW753451
35	47.8	21.1	441	11	W41083
36	47.8	21.1	470	10	AI197242
37	46.4	20.4	796	11	BI157217
38	46	20.3	703	11	BI157489
39	45.2	19.9	959	11	BF100527
40	44	19.4	178	10	AW844444
41	43.6	19.2	496	10	AW753448
42	43	18.9	515	13	AO526647
43	39.8	17.5	128	10	AM603469
44	38.2	16.8	364	10	AM284592
45	38.2	16.8	630	11	BG051390

ALIGNMENTS

RESULT 1
LOCUS: BF083321
DEFINITION RCL-CT0249-130900-212-g08 CT0249 Homo sapiens cDNA, mRNA sequence.
VERSION BF083321
KEYWORDS EST.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 455)
Dias Neto E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunschein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ct2-RCL-CT0249-130900-212-g08&ts=2000-09-13&tt=1)
Seq primer: puc 18 forward

High quality sequence start: 18
High quality sequence stop: 455.
Location/Qualifiers

FEATURES
Source

1. 455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="CT0249"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site-1: SmaI; Site-2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 131 a 101 c 117 g 106 t
ORIGIN

Query Match 100.0%; Score 227; DB 11; Length 455;
Best Local Similarity 100.0%; Pred. No. 1e-56;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTGCGGCTCTGGAGAGTTAAGAGCCAGACGAGAGTATACCCAGAGAGTGA 60
DB 121 GTGCGGCTCTGGAGAGTTAAGAGCCAGACGAGAGTATACCCAGAGAGTGA 180
OY 61 GCACTGTACATACCTGGCTGGATTGAGATGATGAATACATGATCCAGACCAAGCT 120
DB 181 GCACTGTACATACCTGGCTGGATTGAGATGATGAATACATGATCCAGACCAAGCT 240
OY 121 GAAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 241 GAAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
OY 181 GGCTCATTTTGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 227
DB 301 GGCTCATTTTGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 347

RESULT 2
BE927204 598 bp mRNA EST 02-OCT-2000
LOCUS RCI-CT0249-240800-211-h10 CT0249 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE927204
ACCESSION BE927204.1 GI:10453280
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
CONTACT: Simpson A.J.G.
LABORATORY OF Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?rl=ct2-RC1-CT0249-240

800-211-h10ct3-2000-08-24ct4-1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 598.
Location/Qualifiers

FEATURES
Source

1. 598
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="CT0249"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site-1: SmaI; Site-2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 150 a 133 c 130 g 185 t
ORIGIN

Query Match 100.0%; Score 227; DB 11; Length 598;
Best Local Similarity 100.0%; Pred. No. 1e-56;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GTGCGGCTCTGGAGAGTTAAGAGCCAGACGAGAGTATACCCAGAGAGTGA 60
DB 531 GTGCGGCTCTGGAGAGTTAAGAGCCAGACGAGAGTATACCCAGAGAGTGA 492
OY 61 GCACTGTACATACCTGGCTGGATTGAGATGATGAATACATGATCCAGACCAAGCT 120
DB 491 GCACTGTACATACCTGGCTGGATTGAGATGATGAATACATGATCCAGACCAAGCT 432
OY 121 GAAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 431 GAAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
OY 131 GGCTCATTTTGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 227
DB 371 GGCTCATTTTGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 325

RESULT 3
BF083270 528 bp mRNA EST 18-OCT-2000
LOCUS RCI-CT0249-110900-214-b01 CT0249 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF083270
ACCESSION BF083270.1 GI:10871100
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
CONTACT: Simpson A.J.G.
LABORATORY OF Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL
<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC1-CT0249-110>
900-214-b01et3=2000-09-11et4=1)
Seq primer: puc 18 forward
High quality sequence start: 33
High quality sequence stop: 526.
Location/Qualifiers

FEATURES

source

1..528

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="CT0249"

/dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

158 a 113 c 127 g 130 t

ORIGIN

Query Match 99.3%; Score 225.4; DB 11; Length 528;
Best Local Similarity 99.6%; Pred. No. 3.1e-56;
Matches 226; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGCGGCTCTGGGAGAGTTACGACGACGAGAGTGTACCCGACAGAGGA 60
DB 108 GGGCGGCTCTGGGAGAGTTACGACGACGAGAGTGTACCCGACAGAGGA 167
OY 61 GCACCTACATACCTGGCTGGATTGAGAAATGATCAATGATCCACCAAGCT 120
DB 168 GCACCTACATACCTGGCTGGATTGAGAAATGATCAATGATCCACCAAGCT 227
OY 121 GAATTAATTAAGAGTATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCTCGGA 180
DB 228 GAATTAATTAAGAGTATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCTCGGA 287
OY 181 GGCTCATTTGTGGCTGTGATGCCAAATGTCCTACCTGATCT 227
DB 288 GGCTCATTTGTGGCTGTGATGCCAAATGTCCTACCTGATCT 334

RESULT 4

BF083396/c

LOCUS BF083396 315 bp mRNA EST 18-OCT-2000
DEFINITION RC6-CT0437-120900-011-B12 CT0437 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF083396
VERSION BF083396.1 GI:10877226
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

AUTHORS 1 (bases 1 to 315)
Dias Neto, E., Garcia Correia, R., Verjovski-Almeida, S., Bioness, M., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costra, F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brustein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC6-CT0437-120>
900-011-B12et3=2000-09-12et4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 315.
Location/Qualifiers

FEATURES

source

1..315

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="CT0437"

/dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

78 a 73 c 72 g 92 t

ORIGIN

Query Match 74.2%; Score 168.4; DB 11; Length 315;
Best Local Similarity 99.4%; Pred. No. 1.9e-39;
Matches 169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 128 GGAGCACTGTACATACCTGGCTGGATTGAGAAATGATCAATGATCCACCAAGA 117
DB 315 GAAGCACTGTACATACCTGGCTGGATTGAGAAATGATCAATGATCCACCAAGA 256
OY 158 CCGAATTAATTAAGAGTATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCTCG 177
DB 255 CCGAATTAATTAAGAGTATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCTCG 196
OY 178 GGAGCACTGTGGCTGTGATGCCAAATGTCCTACCTGATCT 227
DB 295 GGAGCACTGTGGCTGTGATGCCAAATGTCCTACCTGATCT 146

RESULT 5

A1660957/c

LOCUS A1660957 734 bp mRNA EST 18-DEC-1999
DEFINITION wF020408.X1 Soares, Dieckgraefe_colon_NHUC Homo sapiens cDNA clone
IMAGE:2351151 3' similar to TR:08826 08826 GDB-5 PROTEIN.;, mRNA
sequence.
ACCESSION A1660957
VERSION A1660957.1 GI:4764540
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

AUTHORS 1 (bases 1 to 734)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCI-CGAP <http://www.ncbi.nlm.nih.gov/cgi/gap>.
National Cancer Institute, Cancer genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL

Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
This clone is available royalty-free through LNCI; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 820 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 477.
Location/Qualifiers

FEATURES

source

1..734

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="IMAGE:2351151"

/note="Soares, Dieckgraefe_colon_NHUC"

Query Match	45.5%	Score 103.2	DB 10	Length 469
Best Local Similarity	72.6%	Pred. No. 3.7e-20		
Matches 164	Conservative	0	Mismatches 53	Indels 9
				Gaps 2

Imotani, K., Ishii, Y., Itoh, M., Izeva, M., Kato, H., Kawai, J.,
Kojima, Y., Kono, H., Kouda, M., Koyas, S., Kurihara, C., Matsuyma,
Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki,
Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, R., Sano,
Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T.,

5 GGGCTCTGGGAGAGTTAACGCAGCCAGAGAGTGATACCCAGCAGAGTGGAGCAC 64

51 GCACGTGACATACCTGG -GTGATTGAGAAATGATGAATCAATGGAATCCAC 112
 55 GCACCTGACATACCTGGATCGATTGAGAAATGATGAATCAATGGAATCCAC 3
 61 GTCCGGGCGCTGGAGAGAGATTAAACCCACGACGAGAGATGATACCCGACGAGATGGA 60
 115 GTGGGGGCTCTGGAGAGATTAAACCCACGACGAGAGATGATACCCGACGAGATGGA 56

RESULT 10
LOCUS BG220001 794 bp mRNA EST 21-APR-2001
DEFINITION RST36685 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG220001
VERSION BG220001.1 GI:13746022
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 794)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Dahl,T., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Krashoc,D., McElligott,K., Clark,S., Mays,R., Smith,E.,
Veloso,N., Hess,J., Colhren,K., Lo,K., Offenbacher,J., Danzig,J.
and Ducar,M.
TITLE Creation of Genome-wide Protein Expression Libraries using Random
Activation of Gene Expression
JOURNAL Nat. Biotechnol. 19 (5), 440 (2001) In press
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 328.
FEATURES
source
1..794
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="Hr1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is Hr1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in Hr1080 under normal circumstances."
BASE COUNT 267 a 162 c 159 g 206 t
ORIGIN
Query Match 41.8%; Score 94.8; DB 11; Length 794;
Best Local Similarity 66.8%; Pred. No.1.3e-17;
Matches 151; Conservative 0; Mismatches 72; Indels 3; Gaps 1;
QY 1 GTGCGGCTCTGGAGAGTTAAAGCAGCAGAGAGTGAATACCCAGCAGAGTGA 60
DB 80 GTTCGGCTCATGAGAGAGCAAACTGCGAGCTAAATTAAGGCTCCACGTAATAGA 139
QY 61 GCACTGTACTACTGGCTGGATTTGAGATGATGAATTAATGAATCCACCAAGCT 120
DB 140 GCCCGCTACATACAGGCTGGTGAAGGAGGAAATTAAGCAAAACCGCAAGACT 199
QY 121 GAATTAATTAAGATGATGTTCAACACAGCAAGTGTGTTCAAGCAGACATCTCGGA 180
DB 200 GAATTTGAT---GAGGATATCTAGACCACTTGAGAGATTTGACCCACCAAGCTCCGGA 256
QY 181 GGCTCATTTGGCTTCTGATGTCCCAATGCTCCATACCTGATC 226
DB 257 GGTGCATTTGTGTATCAACAAGTCCCAAGCTTCCCTGCTGACC 302
RESULT 11
LOCUS BG216984 813 bp mRNA EST 21-APR-2001
DEFINITION RST36685 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG216984
VERSION BG216984.1 GI:13743005
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 813)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Dahl,T., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Krashoc,D., McElligott,K., Clark,S., Mays,R., Smith,E.,
Veloso,N., Hess,J., Colhren,K., Lo,K., Offenbacher,J., Danzig,J.
and Ducar,M.
TITLE Creation of Genome-wide Protein Expression Libraries using Random
Activation of Gene Expression
JOURNAL Nat. Biotechnol. 19 (5), 440 (2001) In press
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 481.
FEATURES
source
1..813
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="Hr1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is Hr1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in Hr1080 under normal circumstances."
BASE COUNT 266 a 173 c 160 g 214 t
ORIGIN
Query Match 41.8%; Score 94.8; DB 11; Length 813;
Best Local Similarity 66.8%; Pred. No.1.3e-17;
Matches 151; Conservative 0; Mismatches 72; Indels 3; Gaps 1;
QY 1 GTGCGGCTCTGGAGAGTTAAAGCAGCAGAGAGTGAATACCCAGCAGAGTGA 60
DB 80 GTTCGGCTCATGAGAGAGCAAACTGCGAGCTAAATTAAGGCTCCACGTAATAGA 139
QY 61 GCACTGTACTACTGGCTGGATTTGAGATGATGAATTAATGAATCCACCAAGCT 120
DB 140 GCCCGCTACATACAGGCTGGTGAAGGAGGAAATTAAGCAAAACCGCAAGACT 199
QY 121 GAATTAATTAAGATGATGTTCAACACAGCAAGTGTGTTCAAGCAGACATCTCGGA 180
DB 200 GAATTTGAT---GAGGATATCTAGACCACTTGAGAGATTTGACCCACCAAGCTCCGGA 256
QY 181 GGCTCATTTGGCTTCTGATGTCCCAATGCTCCATACCTGATC 226
DB 257 GGTGCATTTGTGTATCAACAAGTCCCAAGCTTCCCTGCTGACC 302
RESULT 12
LOCUS BG965506 528 bp mRNA EST 12-JUN-2001
DEFINITION 602830872F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4985705 5',
mRNA sequence.
ACCESSION BG965506
VERSION BG965506.1 GI:14353143
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 528)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
unpublished (1999)

Eukaryota: Eutelezoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Mus: 1 (bases 1 to 527)

Marta, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Pearson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Rittenberg, E., Korn, S., Shin, T., Jackson, J., Cardenas, M., McCann, R.,

```

07 181 GGCCTCATTTGGCTGCTGATGTCGCCAAATGCTCCATTACCTATC 228
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
08 264 GGTCCCTTTGGATGCTCAATGTCATGTCATGCTGCTATC 309
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Search completed: April 3, 2002, 20:27:41
Job time: 40224 sec

```

Search completed: April 3, 2002, 20:27:41
Job time: 40224 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:54:16 ; Search time 521.98 seconds
(without alignments)
98.491 Million cell updates/sec

Title: US-09-049-696-13

Perfect score: 227
Sequence: 1 GTCGCGCTCTGCGAGAGT.....AATGCTCCACTGATCT 227

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query #	Match Length	DB ID	Description
1	227	100.0	878	1	US-08-469-667-8
2	227	100.0	878	5	PCT-US95-07289-8
3	88.6	39.0	576	4	US-09-385-982-23
4	80.6	35.5	611	4	US-09-385-982-27
5	67.8	29.9	742	4	US-09-385-982-33
6	66.8	29.4	595	4	US-09-385-982-25
7	28.2	12.4	868	3	US-08-961-083-19
8	28.2	12.4	6371	3	US-08-836-325-13
9	28.2	12.4	6404	3	US-08-836-325-14
10	27.8	12.2	35408	4	US-08-973-334-3
11	27.8	12.2	35408	4	US-09-563-869A-3
12	27.8	12.2	35408	4	US-08-549-489-3
13	27.6	12.2	10754	2	US-08-966-958-1
14	27.6	12.2	10754	2	US-09-215-817-1
15	27.6	12.2	10754	2	US-09-342-353-1
16	27.6	12.2	815	3	US-08-433-263B-1
17	27.2	12.0	1578	3	US-08-504-878A-1
18	27.2	12.0	3381	3	US-08-937-195-1
19	27.2	12.0	3381	3	US-08-937-195-2
20	27.2	12.0	3381	4	US-08-915-152-1
21	27.2	12.0	3381	4	US-08-915-152-2
22	27.2	12.0	3381	5	PCT-US96-07627-1
23	27.2	12.0	3381	5	PCT-US96-07627-2
24	26.8	11.8	1689	1	US-07-991-867B-41
25	26.8	11.8	1689	2	US-08-544-332-41
26	26.8	11.8	3120	1	US-08-222-616-22
27	26.8	11.8	3120	5	PCT-US95-04228-22

28	26.8	11.8	3475	1	US-08-222-299-3	Sequence 3, Appl1
29	26.8	11.8	3475	2	US-08-434-878-3	Sequence 3, Appl1
30	26.8	11.8	3475	5	PCT-US95-03718-3	Sequence 3, Appl1
31	26.8	11.8	3476	1	US-08-183-211-1	Sequence 1, Appl1
32	26.8	11.8	3476	5	PCT-US95-00176A-1	Sequence 1, Appl1
33	26.8	11.8	3501	1	US-07-977-451-3	Sequence 3, Appl1
34	26.8	11.8	3501	1	US-08-252-517-3	Sequence 3, Appl1
35	26.8	11.8	3501	1	US-07-906-397A-3	Sequence 3, Appl1
36	26.8	11.8	3501	1	US-08-601-891-3	Sequence 3, Appl1
37	26.8	11.8	3501	2	US-09-021-324-3	Sequence 3, Appl1
38	26.8	11.8	3501	5	PCT-US92-05401-3	Sequence 3, Appl1
39	26.8	11.8	3501	5	PCT-US92-09893-3	Sequence 3, Appl1
40	26.8	11.8	4071	4	US-09-091-117-3	Sequence 3, Appl1
41	26.8	11.8	8457	1	US-07-991-867B-1	Sequence 1, Appl1
42	26.8	11.8	8457	2	US-08-544-332-1	Sequence 1, Appl1
43	26.8	11.8	14636	2	US-08-715-131-1	Sequence 6, Appl1
44	26.6	11.7	1330	4	US-08-715-131-1	Sequence 1, Appl1
45	26.6	11.7	1330	4	US-09-221-753-1	Sequence 1, Appl1

ALIGNMENTS

RESULT: 1

US-08-469-667-8
Sequence 8, Application US/08469667

Patent No. 5733748

GENERAL INFORMATION:

APPLICANT: Yu, Guo-Liang

APPLICANT: Rosen, Craig

TITLE OF INVENTION: Colon Specific Genes and Proteins

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carella, Byrne, Bain, Gillfillan, Cecchi,

ADDRESSER: Stewart & Olstein

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: NJ

COUNTRY: USA

ZIP: 07068-1739

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,667

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Perraro, Gregory D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-435

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 878 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2..685

Query/Match 100.0%; Score 227; DB 1; Length 878;

Best Local Similarity 100.0%; Pred. No. 3e-71;

Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

QY 1 GTCCGGGCTCTGGAGAGTTAAGCAGCCAGAGAGTATACCCAGAGAGTGA 60
    |||||||
DB 56 GTCCGGGCTCTGGAGAGTTAAGCAGCCAGAGAGTATACCCAGAGAGTGA 115
QY 61 GCACGTACATACCTGGCTGGATTGAGATGAAATACATGGAATCCACCAAGACT 120
    |||||||
DB 116 GCACGTACATACCTGGCTGGATTGAGATGAAATACATGGAATCCACCAAGACT 175
QY 121 GAAATTAATAGATGATGTTCAACACAGCAAGAGTGTTCAGAGAACATCTCGGGA 180
    |||||||
DB 176 GAAATTAATAGATGATGTTCAACACAGCAAGAGTGTTCAGAGAACATCTCGGGA 235
QY 181 GGCTCATTTGTGGCTTCTGATGTCGCCAATGCTCCATACCTATCT 227
    |||||||
DB 236 GGCTCATTTGTGGCTTCTGATGTCGCCAATGCTCCATACCTATCT 282

```

RESULT 2
PCT-US95-07289-8
Sequence 8, Application PC/TUS9507289

```

GENERAL INFORMATION:
APPLICANT: Yu, Guo-Liang
APPLICANT: Rosen, Craig
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Giffillan, Cecchi,
ADDRESSEE: Stewart & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07289
FILING DATE: 06-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-265
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
PCT-US95-07289-8

```

Query Match 100.0%; Score 227; DB 5; Length 878;
Best Local Similarity 100.0%; Pred. No. 3e-71;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTCCGGGCTCTGGAGAGTTAAGCAGCCAGAGAGTATACCCAGAGAGTGA 60
    |||||||
DB 56 GTCCGGGCTCTGGAGAGTTAAGCAGCCAGAGAGTATACCCAGAGAGTGA 115
QY 61 GCACGTACATACCTGGCTGGATTGAGATGAAATACATGGAATCCACCAAGACT 120
    |||||||
DB 116 GCACGTACATACCTGGCTGGATTGAGATGAAATACATGGAATCCACCAAGACT 175

```

```

QY 121 GAAATTAATAGATGATGTTCAACACAGCAAGAGTGTTCAGAGAACATCTCGGGA 180
    |||||||
DB 176 GAAATTAATAGATGATGTTCAACACAGCAAGAGTGTTCAGAGAACATCTCGGGA 235
QY 181 GGCTCATTTGTGGCTTCTGATGTCGCCAATGCTCCATACCTATCT 227
    |||||||
DB 236 GGCTCATTTGTGGCTTCTGATGTCGCCAATGCTCCATACCTATCT 282

```

RESULT 3
US-09-385-982-23
Sequence 23, Application US/09385982

```

GENERAL INFORMATION:
PATENT NO. 6262334
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLES OF INVENTION: PRODUCTS: II
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
EARLIER FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(576)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-23

```

Query Match 39.0%; Score 88.6; DB 4; Length 576;
Best Local Similarity 66.5%; Pred. No. 4e-22;
Matches 143; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

```

QY 1 GTCCGGGCTCTGGAGAGTTAAGCAGCCAGAGAGTATACCCAGAGAGTGA 60
    |||||||
DB 167 gtccgggctctggagagtttaagcagccagagagtataccagagagtgga 226
QY 61 GCACGTACATACCTGGCTGGATTGAGATGAAATACATGGAATCCACCAAGACT 120
    |||||||
DB 227 gccgcgtacataccagcctgggtgagtgagcaaggaattgaagcaaccgccagagact 286
QY 121 GAAATTAATAGATGATGTTCAACACAGCAAGAGTGTTCAGAGAACATCTCGGGA 180
    |||||||
DB 267 gaattgat--gaggtactcagaccacttggagagattcagccgaagacatccgga 343
QY 131 GGCTCATTTGTGGCTTCTGATGTCGCCAATGCTCC 215
    |||||||
DB 344 ggtgcatltgtgtatcacacaagltcccaagccttcc 378

```

RESULT 4

US-09-345-982-27
Sequence 27, Application US/09385982

```

GENERAL INFORMATION:
PATENT NO. 6262334
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLES OF INVENTION: PRODUCTS: II
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111

```


EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 27
LENGTH: 611
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(611)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-27

Query Match 35.5%; Score 80.6; DB 4; Length 611;
Best Local Similarity 65.6%; Pred. No. 2.9e-19;
Matches 149; Conservative 0; Mismatches 74; Indels 4; Gaps 2;

QY 1 GTCCGGGCTCTGGAGAGTTAAGCAGCAGGAGAGTGAATACCCAGCAGAGTGA 60
DB 168 gtccgggctcctagagagaaacacacgccaagctaaattacgacctccatgaataga 227
QY 61 GCACGTACATACCTGCTGATGATGAGATGAATACATGATCCACCAAGACT 120
DB 228 gcacgtactactacgaagctgagtagaagcgggaaattgaagcaaacccgccaagcct 287
QY 121 GAATTAATTAAGATGATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCCTCGGA 180
DB 288 gaattgat--gagagatactcagaccaccttgagagattcagcgaaacagatcccca 344
QY 181 GGTCAATTTGTGGCTTC-TGATGCCCAATGCTCCCATACCGATC 226
DB 345 ggtcatttgtgtgtatcacaagtcacaaccttcctcctcctgacc 391

RESULT 5

US-09-385-982-33
Sequence 33, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS: II
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 742
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(742)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-33

Query Match 29.9%; Score 67.8; DB 4; Length 742;
Best Local Similarity 68.6%; Pred. No. 1.2e-14;
Matches 109; Conservative 0; Mismatches 47; Indels 3; Gaps 1;

QY 168 ACATACCTGCTGATGATGAGATGATGAATACATGGAATCCACCAAGACTGAATTA 127
DB 168 acataccagctggtgtagtgaacggaatltgaagaaacccgccaagacctgaattg 60
QY 168 ATAGATGATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCCTCGGAGAGCTCAT 187
DB 168 atagagatgac--tcaagcacaccttgagagattcagcgaaacagatccgagagtgcat 117
QY 188 TTGTGCTCTGATGCCCAATGCTCCCATACCTGATC 226
DB 188 ttgtgtatcacaagtcacaagctcccttccttgctgacc 156

RESULT 6

US-09-385-982-25
Sequence 25, Application US/09385982
Patent No. 6262334

GENERAL INFORMATION:

APPLICANT: ENDEGE, WILSON O., ET AL.

TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

TITLE OF INVENTION: PRODUCTS: II

FILE REFERENCE: CCDNA-260XX

CURRENT APPLICATION NUMBER: US/09/385,982

CURRENT FILING DATE: 1999-08-30

EARLIER APPLICATION NUMBER: 09/328,111

EARLIER FILING DATE: 1999-06-08

EARLIER APPLICATION NUMBER: 60/117,393

EARLIER FILING DATE: 1999-01-27

EARLIER APPLICATION NUMBER: 60/098,639

EARLIER FILING DATE: 1998-08-31

NUMBER OF SEQ ID NOS: 544

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 25

LENGTH: 595

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(595)

OTHER INFORMATION: n = A,T,C or G

US-09-385-982-25

Query Match

29.4%; Score 66.8; DB 4; Length 595;

Best Local Similarity 67.9%; Pred. No. 2.3e-14;

Matches 108; Conservative 0; Mismatches 48; Indels 3; Gaps 1;

QY 168 ACATACCTGCTGATGATGAGATGATGAATACATGGAATCCACCAAGACTGAATTA 127
DB 168 acataccagctggtgtagtgaacggaatltgaagaaacccgccaagacctgaattg 60
QY 128 ATAGATGATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCCTCGGAGAGCTCAT 187
DB 128 atagagatgac--tcaagcacaccttgagagattcagcgaaacagatccgagagtgcat 117
QY 188 TTGTGCTCTGATGCCCAATGCTCCCATACCTGATC 226
DB 188 ttgtgtatcacaagtcacaagctcccttccttgctgacc 156

RESULT 7

US-08-981-083-19

Sequence 19, Application US/08961083

Patent No. 6159469

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 868 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-19

Query Match 12.4%; Score 28.2; DB 3; Length 868;
Best Local Similarity 50.4%; Pred. No. 1.5; Mismatches 69; Conservative 0; Indels 0; Gaps 0;

QY 82 ATTGAGATGATGAATATACATGAAATCCACCAAGCTGAAATTAATAGATGATGT 141
DB 470 ACGATATAGCTTAGACAACTTGATTAAGAAAGTAAGATTAATTAATGATCCCTGCT 529
QY 142 CAACACAGCAAGATGTTTTCAGCAGACATCTCTGGAGGCTCATTTGGCTGTGAT 201
DB 530 GAAAGAAAGTCAATGTGTACAGCAGGAGGACATTCATTAATCTTCTTAAGCTTATGCT 589
QY 202 GTCCCAATGCTCCCAT 218
DB 590 GTCCCAAGTCTTACAT 606

RESULT 8
US-08-836-325-13
Sequence 13, Application US/08836325
Patent No. 6110672
GENERAL INFORMATION:
APPLICANT: Mandel, Gail
APPLICANT: Halegoua, Simon
TITLE OF INVENTION: Peripheral Nervous System Specific
TITLE OF INVENTION: Sodium Channels, DNA Encoding Therefor, Crystallization,
TITLE OF INVENTION: X-ray Diffraction, Computer Molecular Modeling, Rational
TITLE OF INVENTION: Drug Design, Drug Screening, and Methods of Making and Using
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Ave., N. W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,325
FILING DATE: 2-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/14251
FILING DATE: 02-NOV-1995
PRIOR APPLICATION DATA: 08/482,401
APPLICATION NUMBER: 08/482,401
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/334,029
FILING DATE: 02-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0917.0240002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 6371 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: DNA (genomic)
US-08-836-325-13

Query Match 12.4%; Score 28.2; DB 3; Length 6371;
Best Local Similarity 53.1%; Pred. No. 4.2; Mismatches 60; Conservative 0; Indels 53; Gaps 0;

QY 45 ACCCCAGCAGATGAGCAGCATACATCTGCGATGAGATGATGAATATCAATG 104
DB 3392 ACCCTCTGATTTGATGAGCAGCATGATTAACCTTTGCTGAGAGGAGGAGGAGGAGG 3441
QY 105 GATCCACCAAGCCTGAAATTAATAGATGATGTTCAACACAGCAAGTGT 157
DB 3442 GGTGAAGCTATGATTCGATGAGCCAGAGGCGCTTTCACAGATGTTGTGT 3494

RESULT 9
US-08-836-325-14
Sequence 14, Application US/08836325
Patent No. 6110672
GENERAL INFORMATION:
APPLICANT: Mandel, Gail
APPLICANT: Halegoua, Simon
TITLE OF INVENTION: Peripheral Nervous System Specific
TITLE OF INVENTION: Sodium Channels, DNA Encoding Therefor, Crystallization,
TITLE OF INVENTION: X-ray Diffraction, Computer Molecular Modeling, Rational
TITLE OF INVENTION: Drug Design, Drug Screening, and Methods of Making and Using
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Ave., N. W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,325
FILING DATE: 2-MAY-1997
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/14251
FILING DATE: 02-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/482,401
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/334,029
FILING DATE: 02-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0917.0240002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 6404 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: DNA (genomic)
US-08-836-325-14

Query Match 12.4%; Score 28.2; DB 3; Length 6404;
Best Local Similarity 53.1%; Pred. No. 4.2;
Matches 60; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 45 ACCCCAGCAGAGTGGACACTGTACATCGCTGGATGGAATGATGAATACATG 104
DB 3415 AGCTCCCTCAGAGTGACACAGTGTATACCTTTGGCTGGAGAGGAGGAGCAGAG 3474
QY 105 GAATCCACCAAGACTTAATTTATAGATGATGTTCAACACAGCAGTGT 157
DB 3475 GCTGAACCTTGAATTCGATGACGACGAGGCCCTTTTACAGATGTTGTGT 3527

RESULT 10
US-08-973-334-3
Sequence 3, Application US/08973334
Patent No. 6261551
GENERAL INFORMATION:
APPLICANT: Wilson, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Gao, Guang-Ping
TITLE OF INVENTION: Recombinant Adenovirus and Adeno-
Associated Virus, Cell Lines, and
Methods of Production and Use
TITLE OF INVENTION: Methods of Production and Use
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Box 457, 321 No. 6261551stowm Road
CITY: Spring House
STATE: PA
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release 1.0 Version 1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,334
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/462,014
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/549,489

FILING DATE: 27-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNP012CIPUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 540-9206
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 35408 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: other nucleic acid
US-08-973-334-3

Query Match 12.2%; Score 27.8; DB 4; Length 35408;
Best Local Similarity 54.4%; Pred. No. 14;
Matches 56; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 24 CGCAGCCAGACGAGATGATACCCAGCAGAGTGAGACACTGTACATACCTGCTGAT 83
E3 911 CGCGCGCGCGCGAGCGGATCAGCCAGCGGCGGCGGCGCAATTCGCGGATCGAA 970
QY 34 TGAGATGATGAATTAATGAATCCACCAAGACTGAATTT 126
DB 971 AGACCTCTTAACCAAAAAGAGTACCATGTGCTTACTT 1013

RESULT 11
US-09-563-869A-3
Sequence 3, Application US/09563869A
Patent No. 6270996
GENERAL INFORMATION:
APPLICANT: Wilson, James M.
APPLICANT: Fisher, Krishna J.
APPLICANT: Gao, Guang-Ping
TITLE OF INVENTION: Recombinant Adenovirus and Adeno-
Associated Virus, Cell Lines, and
Methods of Production and Use
TITLE OF INVENTION: Methods of Production and Use
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Box 457, 321 No. 6270996stowm Road
CITY: Spring House
STATE: PA
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release 1.0 Version 1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/563,869A
FILING DATE: 03-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/973,334
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/549,489
FILING DATE: 27-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNP012CIPUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 540-9206
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 35408 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: other nucleic acid
SEQUENCE DESCRIPTION: SEQ ID NO: 3
US-09-563-869A-3

Query Match 12.2%; Score 27.8; DB 4; Length 35408;
Best Local Similarity 54.4%; Pred. No. 14;
Matches 56; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

24 CGCAGCCAGACGAGAGATACCCAGCAGAGTGGAGCATGTACATACCTGGCTGAT 83
DB 911 CGCGGGGGGGGAGGGGATGACGACCGGGTGGCGGCAATTCGGGGGATCGAA 970
QY 84 TGAGATGATGAATACATGCAATCCACCAAGACTGAAAT 126
DB 971 AGAGCTGTAAAGCAAAAAGAGTCCATGTGTTACTT 1013

RESULT 12
US-08-549-489-3
Sequence 3, Application US/08549489
Patent No. 6281010
GENERAL INFORMATION:
APPLICANT: Wilson, James M.
APPLICANT: Gao, Guang-Ping
TITLE OF INVENTION: No. 6281010el Adenovirus Gene Therapy Vehicle
TITLE OF INVENTION: and Cell Line
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Box 457, 321 No. 6281010ristown Road
CITY: Spring House
STATE: PA
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/549,489
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/462,014
FILING DATE: 08-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPVNO13
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 540-9206
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 35408 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: other nucleic acid
US-08-549-489-3

Query Match 12.2%; Score 27.8; DB 4; Length 35408;
Best Local Similarity 54.4%; Pred. No. 14;
Matches 56; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

24 CGCAGCCAGACGAGAGATACCCAGCAGAGTGGAGCATGTACATACCTGGCTGAT 83
DB 911 CGCGGGGGGGGAGGGGATGACGACCGGGTGGCGGCAATTCGGGGGATCGAA 970
QY 84 TGAGATGATGAATACATGCAATCCACCAAGACTGAAAT 126
DB 971 AGAGCTGTAAAGCAAAAAGAGTCCATGTGTTACTT 1013

RESULT 13
US-08-966-958-1
Sequence 1, Application US/08966958
Patent No. 5928908
GENERAL INFORMATION:
APPLICANT: Dunn, John
APPLICANT: Randesi, Matthew
TITLE OF INVENTION: METHODS FOR INTRODUCING UNIDIRECTIONAL
TITLE OF INVENTION: DELETIONS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET: P. O. Box 5000
CITY: Upton
STATE: New York
COUNTRY: US
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/966,958
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: AUI97-14
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 344-3729
TELEFAX: (516) 344-3341
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10754 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-966-958-1

Query Match 12.2%; Score 27.6; DB 2; Length 10754;
Best Local Similarity 56.7%; Pred. No. 9;
Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

100 CATGGAATCCAGCAAGACTGAAATTAATTAAGATGATGTCACACCAAGAGTGT 159
QY 754 CTATTAATCTTACCAATCTTGAATCTGTAAGAAAGTTCATATACGACCAAGATTTT 7623
DB 754 CTATTAATCTTACCAATCTTGAATCTGTAAGAAAGTTCATATACGACCAAGATTTT 7623
QY 150 TTCAGCAGAACATCTCGGAGGCTCATTT 189
DB 7624 TTAAGAAAGCAAGCAAGTAAAT 7653

RESULT 14
US-09-215-817-1
Sequence 1, Application US/09215817
Patent No. 5968786
GENERAL INFORMATION:
APPLICANT: Dunn, John
APPLICANT: Randesi, Matthew
TITLE OF INVENTION: METHODS FOR INTRODUCING UNIDIRECTIONAL

TITLE OF INVENTION: DELETIONS
 NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Brookhaven National Laboratory
 STREET: P.O. Box 5000
 CITY: Upton
 STATE: New York
 COUNTRY: US
 ZIP: 11973
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/215,817
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/966,958
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Bogosian, Margaret
 REGISTRATION NUMBER: 25,324
 REFERENCE/DOCKET NUMBER: AUI97-14
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (516) 344-3341
 TELEFAX: (516) 344-3729
 INFORMATION FOR SEQ. ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10754 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-09-215-817-1

Query Match 12.2%; Score 27.6; DB 2; Length 10754;
 Best Local Similarity 56.7%; Pred. No. 9;
 Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 100 CAATGATCCACCAAGCTGAATTAATAGGATGATGTTCAACACCAAGAGTGTT 159
 DB 7564 CTATTAATCTACCAATCTGAACCTGTAAGAGCTGATATTCAGACCAAGATTTT 7623
 QY 160 TTCAGCAGACATCTCGGAGGCTCATTT 189
 DB 7624 TTAATAAGAACATGCGAGGTAATTT 7653

RESULT 15
 US-09-342-353-1
 Sequence 1, Application US/09342353
 Patent No. 6248569
 GENERAL INFORMATION:
 APPLICANT: Dunn, John
 TITLE OF INVENTION: METHOD FOR INTRODUCING UNIDIRECTIONAL NESTED DELETIONS
 FILE REFERENCE: CIP OF U.S. Application 08/966,958
 CURRENT APPLICATION NUMBER: US/09/342,353
 EARLIER FILING DATE: 1999-06-29
 EARLIER APPLICATION NUMBER: 08/966,958
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: Patentin Ver. 2.0
 SEQ. ID NO. 1
 LENGTH: 10754
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-342-353-1

Query Match 12.2%; Score 27.6; DB 4; Length 10754;

Best Local Similarity 56.7%; Pred. No. 9;
 Matches 51; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 100 CAATGATCCACCAAGCTGAATTAATAGGATGATGTTCAACACCAAGAGTGTT 159
 DB 7564 CTATTAATCTACCAATCTGAACCTGTAAGAGCTGATATTCAGACCAAGATTTT 7623
 QY 160 TTCAGCAGACATCTCGGAGGCTCATTT 189
 DB 7624 TTAATAAGAACATGCGAGGTAATTT 7653
 Search completed: April 3, 2002, 20:54:33
 Job time: 41831 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:47 ; Search time 16681 Seconds
(without alignments)
68.284 Million cell updates/sec

Title: US-09-049-696-17

Perfect score: 106
Sequence: 1 GGCATTGCACATTTTAAAAAT.....AAAATAATCATTCATCCTT 106

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estln:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_est1:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	106	100.0	392	10	A1687981 wa77f06.x
2	106	100.0	450	10	AM009764 wa87b05.x
3	106	100.0	466	11	BG195179 RSM14358
4	106	100.0	490	10	AA508854 n122h10.s
5	106	100.0	504	10	A1802693 w17c01.x
6	106	100.0	561	10	A1802756 w18a04.x
7	106	100.0	712	10	AM009763 AM009763
8	106	100.0	716	10	A1660234 wa68g02.x
9	106	100.0	734	10	A1660957 wf20d08.x
10	104.4	98.5	161	10	A1582072 ar96a11.x
11	103	97.2	255	11	D25727 HUMG04094
12	94.8	89.4	501	10	A1721275 as82h08.x

13	92.8	87.5	335	10	A1721121 as73d08.x
14	85.4	80.6	378	10	AA296955 EST112726
15	50.6	47.7	270	10	BB072582 BB072582
16	49	46.2	237	10	AV369383 AV369383
17	49	46.2	268	10	AV375939 AV375939
18	48	45.3	234	10	AV377609 AV377609
19	48	45.3	252	10	AV371415 AV371415
20	48	45.3	249	10	AV371909 AV371909
21	48	45.3	268	10	AV373431 AV373431
22	47.4	44.7	226	10	BB072141 BB072141
23	47.4	44.7	234	10	AV372464 AV372464
24	47.4	44.7	236	10	AV372363 AV372363
25	47.4	44.7	250	10	AV371940 AV371940
26	47.4	44.7	256	10	AV373943 AV373943
27	47.4	44.7	256	10	AV378819 AV378819
28	47.4	44.7	258	10	AV079760 AV079760
29	47.4	44.7	264	10	AV371007 AV371007
30	47.4	44.7	283	10	AV076160 AV076160
31	47.4	44.7	306	10	AV059168 AV059168
32	47.4	44.7	308	10	AV049139 AV049139
33	47.4	44.7	386	10	AA692521 AA692521
34	47.4	44.7	388	10	AA734161 AA734161
35	47.4	44.7	418	10	AA596289 AA596289
36	47.4	44.7	421	10	AA691586 AA691586
37	47.4	44.7	431	10	AA238284 AA238284
38	47.4	44.7	463	10	AA839323 AA839323
39	47.4	44.7	482	10	AA711228 AA711228
40	47.4	44.7	488	10	AA733415 AA733415
41	47.4	44.7	552	10	AA688953 AA688953
42	47.4	44.7	2915	12	AK007466 Mus muscu
43	47.4	44.7	2933	12	AK008659 Mus muscu
44	46.6	44.0	234	10	AV372473 AV372473
45	46.6	44.0	259	10	AV373600 AV373600

ALIGNMENTS

RESULT 1
LOCUS A1687981/C
DEFINITION wa77f06.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2302211 3' similar to TR:088826 088826 COB-5 PROTEIN. ;, mRNA
SEQUENCE.
ACCESSION A1687981
VERSION A1687981.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 392)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40up from Gibco.

FEATURES

source
1..392
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2302211"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pTR73D-Pac (Pharmacia) with
a modified polylinker; site_1: Not I; site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in

vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1 M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 125 a 69 c 67 g 131 t

Query Match 100.0%; Score 106; DB 10; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.9e-19;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTGACATTTTAAATATGTCAGATGAGAGACAGCTGTCATAGCC 60
|||||
DB 281 GGCATTGACATTTTAAATATGTCAGATGAGAGACAGCTGTCATAGCC 222
OY 61 TAGGCTGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 106
|||||
DB 221 TAGGCTGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 176

RESULT 2
AM009764/c 450 bp mRNA EST 08-MAR-2000
LOCUS W887B05.x1 NCI_CGAP_C03 Homo sapiens cDNA clone IMAGE:2504913 3'
DEFINITION similar to TR:088826 O88826 G08-5 PROTEIN.; mRNA sequence.
ACCESSION AM009764
VERSION AM009764.1 GI:5858542
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 450)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bbrp/image/image.html
Insert length: 587 Std Error: 0.00
Seq primer: -40up from Gibco.

FEATURES
source

1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2504913"
/clone_lib="NCI_CGAP_C03"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/note="Vector: pTY73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from 12 pooled bulk tumor samples and primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pTY73 vector. Library went through one round of
normalization."

BASE COUNT 148 a 80 c 78 g 144 t

Query Match 100.0%; Score 106; DB 10; Length 450;
Best Local Similarity 100.0%; Pred. No. 4.8e-19;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTGACATTTTAAATATGTCAGATGAGAGACAGCTGTCATAGCC 60
|||||
DB 282 GGCATTGACATTTTAAATATGTCAGATGAGAGACAGCTGTCATAGCC 223
OY 61 TAGGCTGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 106
|||||
DB 222 TAGGCTGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 177

RESULT 3
BG195179 466 bp mRNA EST 21-APR-2001
LOCUS R514358 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
DEFINITION BG195179
ACCESSION BG195179
VERSION BG195179.1 GI:13716866
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 466)
Harrington, J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Dahl, T., Thornton, M., Ramchandran, R., Whittington, J.,
Lerner, L., Krashock, D., McElligott, K., Clark, S., Mays, R., Smith, E.,
Veloso, N., Hess, J., Cothren, K., Lo, K., Offenbacher, J., Danzig, D.,
and Ducart, M.
Creation of Genome-wide Protein Expression Libraries using Random
Activation of Gene Expression
Nat. Biotechnol. 19 (5), 440 (2001) In press
JOURNAL Contact: Scott J. Cain
COMMENT Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 466.
Location/Qualifiers
1..466
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

FEATURES
source

1..466
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

BASE COUNT 155 a 92 c 74 g 144 t 1 others

Query Match 100.0%; Score 106; DB 11; Length 466;
Best Local Similarity 100.0%; Pred. No. 4.8e-19;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTGACATTTTAAATATGTCAGATGAGAGACAGCTGTCATAGCC 60
|||||
DB 316 GGCATTGACATTTTAAATATGTCAGATGAGAGACAGCTGTCATAGCC 395
OY 61 TAGGCTGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 106
|||||
DB 336 TAGGCTGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 441

RESULT 4
AA508854/c 490 bp mRNA EST 18-AUG-1997
LOCUS AA508854

DEFINITION n122h10.s1 NCI_CGAP_C04 Homo sapiens CDNA clone IMAGE:968803 3',
mRNA sequence.
ACCESSION AA508854
VERSION AA508854.1 GI:2246357
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 490)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.lnl.gov/bcrp/image/image.html
Insert Length: 856 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 404.
Location/Qualifiers
1..490
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:968803"
/clone_1lb="NCI_CGAP_C04"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/note="Vector: pRT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand CDNA was prepared from pooled colon
tumor tissue, and was then primed with a Not I - oligo(dT)
primer. Double-stranded CDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pRT73 vector.
This library is not normalized. Library constructed by
Bento Soares and M. Fatima Bonaldo."

BASE COUNT 155 a 82 c 89 g 164 t
ORIGIN

Query Match 100.0%; Score 106; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 4.7e-19;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTACATTTTAAATATGTCGAGTAGGAGACTGCAGCTGTAATAGCC 60
|||||
DB 317 GGCATTACATTTTAAATATGTCGAGTAGGAGACTGCAGCTGTAATAGCC 258
OY 61 TAGGGCTGAATTTTGTGCAGATAAATAATATCATTCATCCTT 106
|||||
DB 257 TAGGGCTGAATTTTGTGCAGATAAATAATATCATTCATCCTT 212

RESULT 5
A1802693 504 bp mRNA EST 18-DEC-1999
LOCUS A1802693/C
DEFINITION w17c01.x1 Soares Dieckgraefe.colon_NHUC Homo sapiens CDNA clone
IMAGE:2350848 3' similar to TR:088826 088826 GDB-5 PROTEIN. ; mRNA
sequence.
ACCESSION A1802693
VERSION A1802693.1 GI:5368165
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

DEFINITION Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 504)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Insert Length: 727 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 392.
Location/Qualifiers
1..504
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2350848"
/clone_1lb="Soares-Dieckgraefe.colon_NHUC"
/tissue_type="colonic mucosa from 5 ulcerative colitis
patients"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pRT73D-Pac (Pharmacia) with a
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
strand CDNA was primed with a Not I - oligo(dT) primer [5'
GTGTACCAATCTGAGTGGAGCGGCGGCTACTCTTTTCTTTTCTTTT 3']
double-stranded CDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRT73 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraefe (Washington University,
dieckelm.wustl.edu); colonic mucosa represents a range of
disease involvement from mild cryptitis to severe
ulceration, fibrosis, and degeneration. Library
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 150 a 85 c 103 g 163 t
ORIGIN

Query Match 100.0%; Score 106; DB 10; Length 504;
Best Local Similarity 100.0%; Pred. No. 4.7e-19;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTACATTTTAAATATGTCGAGTAGGAGACTGCAGCTGTAATAGCC 60
|||||
DB 120 GGCATTACATTTTAAATATGTCGAGTAGGAGACTGCAGCTGTAATAGCC 61
OY 61 TAGGGCTGAATTTTGTGCAGATAAATAATATCATTCATCCTT 106
|||||
DB 50 TAGGGCTGAATTTTGTGCAGATAAATAATATCATTCATCCTT 15

RESULT 6
A1802756 561 bp mRNA EST 18-DEC-1999
LOCUS A1802756/C
DEFINITION w18a04.x1 Soares Dieckgraefe.colon_NHUC Homo sapiens CDNA clone
IMAGE:2350926 3' similar to TR:088826 088826 GDB-5 PROTEIN. ; mRNA
sequence.
ACCESSION A1802756
VERSION A1802756
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 561)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the

Db	124	GGCATTCCCATTTTAAAAATATGTGGAGGATAGAGACTGCACCTGTCAATAGCC	65
QY	61	TAGGAGCTGAATTTTGTGCAGATAAATAAATAATCAATTCATCCTT	106
Db	64	TAGGAGCTGAATTTTGTGCAGATAAATAAATAATCAATTCATCCTT	19
RESULT	11		
LOCUS	D25727		
DEFINITION	D25727	255 bp	EST
ACCESSION	D25727	Human colon mucosa	Homo sapiens cDNA clone cm2037 3', mRNA sequence.
VERSION	D25727.1	GI:500422	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 255)		
TITLE	Okubo,K., Yoshii,J., Yokouchi,H., Kameyama,M. and Matsubara,K.		
JOURNAL	Global analysis of gene expression in colon mucosa: a large scale random cDNA sequencing analysis		
COMMENT	Unpublished (1994)		
FEATURES	Contact: Okubo,K., Itoh,K., Yoshii,J., Yokouchi,H. and Matsubara,K.		
SOURCE	Institute for Molecular and Cellular Biology		
ORIGIN	Osaka University		
BASE COUNT	3-1 Yamada-Oka,Suita,Osaka 565,Japan.		
LOCUS	76 a	52 c	39 g
Query Match	97.2%;	Score 103;	DB 11;
Best Local Similarity	97.2%;	Pred. No. 3,4e-18;	Length 255;
Matches 103;	Conservative	0;	Mismatches 3;
		Indels	0;
		Gaps	0;
QY	1	GGCATTACATTTTAAATAATATGTGGAGTGCATAGGAGAACTGCAGCTGCATTAAGCC	60
Db	133	GGCATTACATTTTAAATAATATGTGGAGTGCATAGGAGAACTGCAGCTGCATTAAGCC	192
QY	61	TAGGAGCTGAATTTTGTGCAGATAAATAAATAATCAATTCATCCTT	106
Db	193	TAGGAGCTGAATTTTGTGCAGATAAATAAATAATCAATTCATCCTT	238
RESULT	12		
LOCUS	A1721275/c		
DEFINITION	A1721275	501 bp	EST
ACCESSION	A1721275	Human colon HPIR97	Homo sapiens cDNA clone
VERSION	A1721275	IMAGE:2335263 3'	similar to SW-EGIC_BOVIN P54281 EPITHELIAL
KEYWORDS	CHLORIDE CHANNEL PROTEIN ;	mRNA sequence.	
SOURCE	A1721275.1	GI:5038531	
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 501)		
JOURNAL	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Getzel,G., Jost,S., Kitzman,D., Kucab,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.		
COMMENT	WashU-NCI human EST Project		
FEATURES	Unpublished (1997)		

COMMENT:

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel.: 314 286 1800
Fax: 314 286 1810

Email: estewatson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium ([info@image.lnl.gov](#)) for further information.

Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from glbpco

High quality sequence: sbpc 394.
Location/Qualifiers

SOURCE

1 . 501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2335263"
/clone_lid="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/note="Organ: colon; Vector: pTV73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTGAGCGGGAGCCGCCGCCTTTTTTTTTCCTTTTTCCTTTTTCCTTAATAACAATAAATCATTCATTCTT 106 TAGGGCTGAATTTTGTCAGATAAAATAAATAATCATTCTCTT 106 GTAGGCTACTTTTGTCCTATTAATAACAATAAATCATTCATTCTT 18]

BASE COUNT

158 a 83 c 106 g 154 t Barstead."

ORIGIN

Query Match 89.4%; Score 94.8; DB 10; Length 501;
Best Local Similarity 93.4%; Pred. No. 5.le-16;
Matches 99; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

CY

1 GGCAATTCACATTTTAAAAATTTGGAAGTCGATAGAGAAGAACCTGCATTAACC 60
DBS 123 GCCATTCACATTTTAAAATTTATGTGGAAGTCGATAGAGAAGAACCTGCATTAACC 64

ID

61 TAGGGCTGAATTTTGTCAGATAAAATAAATAATCATTCTCTT 106
63 TAGGCTACTTTTGTCCTATTAATAACAATAAATCATTCATTCTT 18

RESULT #3

A1721121 335 bp mRNA EST 10-JUN-1999
LOCUS A1721121.C as73d08.x1 Barstead colon HPLRB7 Homo sapiens CDNA clone
DEFINITION IMAGE:2334351 3' similar to TR:088826 088826 GOB-5 PROTEIN.; ; mrna
sequence.

ACCESSION

A1721121
VERSION A1721121.1 GI:5038377
KEYWORDS EST.
ORGANISM

Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 335)
Hitler.L., Allen.M., Bowles.L., Dubouque.T., Geisel.G., Jost.S.,
Kilzman.D., Kucaba.T., Lacy.M., Le.N., Lennon.G., Marra.J., Martin
.J., Moore.B., Schellenberg.K., Stepcone.M., Tan.F., Theising,B.,
White.Y., Wylie,T., Waterston.R. and Wilson,R.
WashU-NCI Human EST Project
Unpublished (1997)
Contact: Wilison RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the

JOURNAL

TITLE

REFERENCE

AUTHORS

COMMENT

OY 1 GGCATTACATTTTAAATATGTCGAGTGCAGTGCAGTGCATATGACC 60
DB 717 GGCATTACATTTTAAATATGTCGAGTGCAGTGCAGTGCATATGACC 776
OY 61 TAGGGCTGAATTTTGTGCATTAATAA 89
DB 777 TAGGGCTGAATTTTGTGCATTAATAA 805

RESULT 2

PCT-US95-07289-8
Sequence 8, Application PC/TUS9507289
GENERAL INFORMATION:
APPLICANT: Yu, Guo-Liang
APPLICANT: Rosen, Craig
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
STREET: Stewart & Olstein
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07289
FILING DATE: 06-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-265
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
PCT-US95-07289-8

Query Match 72.5%; Score 76.8; DB 5; Length 878;
Best Local Similarity 91.0%; Pred. No. 1.4e-16;
Matches 81; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 GGCATTACATTTTAAATATGTCGAGTGCAGTGCAGTGCATATGACC 60
DB 717 GGCATTACATTTTAAATATGTCGAGTGCAGTGCAGTGCATATGACC 776
OY 61 TAGGGCTGAATTTTGTGCATTAATAA 89
DB 777 TAGGGCTGAATTTTGTGCATTAATAA 805

RESULT 3
US-08-454-557C-50/C
Sequence 50, Application US/08454557C
Patent No. 583670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-50

Query Match 31.3%; Score 33.2; DB 2; Length 2520;
Best Local Similarity 57.8%; Pred. No. 0.04;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 2 GGCATTACATTTTAAATATGTCGAGTGCAGTGCAGTGCATATGACC 61
DB 2412 GACACATGCTTTAATATGTCGAGTGCAGTGCAGTGCATATGACC 2353
OY 62 AGGGCTGAATTTTGTGCATTAATAAATTAATCATTCATC 103
DB 2352 AAAAAGAAATCCTTGATTAATTAATTAACAAAGCATGATC 2311

RESULT 4
US-08-340-426D-50/C
Sequence 50, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D

FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-50

Query Match 31.3%; Score 33.2; DB 2; Length 2520;
Best Local Similarity 57.8%; Pred. No. 0.04;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 2 GCATTCACATTTTAAATATGTGAGAGTGAAGTGCAGCTGTCAATAGCCT 61
DB 2412 GAACACATGCTTTAAATATGACAGTGGAGGAGGCTGATTACACGTCACAAAGCTA 2353
QY 62 AGGCGTGAATTTTGTGCAGTAATAATAATTAATTCATTCATC 103
DB 2352 AAAAAAGATCCTTGATGTGATTATTAATTAACAAAGCATGATC 2311

RESULT 5
US-08-450-673C-50/C
Sequence 50, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-50

Query Match 31.3%; Score 33.2; DB 2; Length 2520;
Best Local Similarity 57.8%; Pred. No. 0.04;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 2 GCATTCACATTTTAAATATGTGAGAGTGAAGTGCAGCTGTCAATAGCCT 61
DB 2412 GAACACATGCTTTAAATATGACAGTGGAGGAGGCTGATTACACGTCACAAAGCTA 2353
QY 62 AGGCGTGAATTTTGTGCAGTAATAATAATTAATTCATTCATC 103
DB 2352 AAAAAAGATCCTTGATGTGATTATTAATTAACAAAGCATGATC 2311

RESULT 6
PCT-US95-17111A-50/C
Sequence 50, Application PC/TUS9517111A
GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-50

Query Match 31.3%; Score 33.2; DB 5; Length 2520;
Best Local Similarity 57.8%; Pred. No. 0.04;
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 2 GCATTCACATTTTAAATATGTGAGAGTGAAGTGCAGCTGTCAATAGCCT 61
DB 2412 GAACACATGCTTTAAATATGACAGTGGAGGAGGCTGATTACACGTCACAAAGCTA 2353
QY 62 AGGCGTGAATTTTGTGCAGTAATAATAATTAATTCATTCATC 103
DB 2352 AAAAAAGATCCTTGATGTGATTATTAATTAACAAAGCATGATC 2311

RESULT 7
US-09-145-287-5/C
Sequence 5, Application US/09125287B

Patent No. 6114602
GENERAL INFORMATION:
APPLICANT: BARG, Rivka
TITLE OF INVENTION: METHOD FOR THE INTRODUCTION OF GENETIC PARTHENOCAPI IN
TITLE OF INVENTION: PLANTS
FILE REFERENCE: INTR. GENETIC PARTHENOCAPI IN PLANTS
CURRENT FILING DATE: 1998-11-09
EARLIER APPLICATION NUMBER: US/09/125,287B
EARLIER FILING DATE: 1997-02-13
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentlin Ver. 2.0
SEQ ID NO 5
LENGTH: 2293
TYPE: DNA
ORGANISM: ROLB GENE
US-09-125-287-5

Query Match 26.4%; Score 28; DB 3; Length 2293;
Best Local Similarity 56.5%; Pred. No. 2;
Matches 52; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 5 TTCACATTTTAAATTTATGTCAGATGAGAACTGCAGCTGCATAGCCTAGG 64
DB 1770 TTACGTCACAAATAGTTGTTTAAAGTGGTATTAATTCATTTTAAAGGAAAT 1711
QY 65 GCTGAATTTTGTGTCAGATTAATTAATTAATC 96
DB 1710 TGAATTTTACGAGAGAAACAAATTTGAAC 1679

RESULT 8
US-09-125-287-1/c
Sequence 1, Application US/09125287B
Patent No. 6114602
GENERAL INFORMATION:
APPLICANT: BARG, Rivka
APPLICANT: SALTS, Tehlan
TITLE OF INVENTION: METHOD FOR THE INTRODUCTION OF GENETIC PARTHENOCAPI IN
TITLE OF INVENTION: PLANTS
FILE REFERENCE: INTR. GENETIC PARTHENOCAPI IN PLANTS
CURRENT FILING DATE: 1998-11-09
EARLIER APPLICATION NUMBER: US/09/125,287B
EARLIER FILING DATE: 1997-02-13
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentlin Ver. 2.0
SEQ ID NO 1
LENGTH: 12839
TYPE: DNA
ORGANISM: TRP-F1 GENOMIC CLONE
FEATURE:
NAME/KEY: unsure
LOCATION: (5)..(11322)
OTHER INFORMATION: "n"s are any nucleic residue
US-09-125-287-1

Query Match 26.4%; Score 28; DB 3; Length 12839;
Best Local Similarity 56.5%; Pred. No. 3.2;
Matches 52; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 5 TTCACATTTTAAATTTATGTCAGATGAGAACTGCAGCTGCATAGCCTAGG 64
DB 10709 TTACGTCACAAATAGTTGTTTAAAGTGGTATTAATTCATTTTAAAGGAAAT 10650
QY 65 GCTGAATTTTGTGTCAGATTAATTAATTAATC 96
DB 10649 TGAATTTTACGAGAGAAACAAATTTGAAC 10618

RESULT 9
US-08-700-626-2
Sequence 2, Application US/08700626
Patent No. 5734038
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN DBI/ACBP-LIKE PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,626
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0115 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1123 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
LIBRARY: PGANMOT01
CLONE: 620984
US-08-700-626-2

Query Match 25.8%; Score 27.4; DB 1; Length 1123;
Best Local Similarity 65.6%; Pred. No. 2.5;
Matches 40; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 33 GATGAGAGAACTGCAGCTGTCATATGAGGCTGAATTTTGTGACATTAATTAATA 92
DB 973 GACTGGAAGAACTGCAGCTGTCATATGAGGCTTCATTAATGAAGAAACATACAAA 1032
QY 93 A 93
DB 1033 A 1033

RESULT 10
US-09-201-641-5
Sequence 5, Application US/09201641A
Patent No. 6232530
GENERAL INFORMATION:
APPLICANT: Cunningham JR, Francis X
APPLICANT: DellaPenna, Dean
TITLE OF INVENTION: Method for Regulating Carotenoid Biosynthesis in
TITLE OF INVENTION: Mariogolds
FILE REFERENCE: Quest 41-162
CURRENT APPLICATION NUMBER: US/09/201,641A
CURRENT FILING DATE: 1998-11-30
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentlin Ver. 2.0
SEQ ID NO 5

OY 3 CATTGACATTTTAAATAATATGTGAGTAGAGACTGCAGCTGTCAATAGCCTA 62
DB 8914 CTTTATATTTTAAAAAATTTGCTACAGTGATCCACAGAGTTCAACCCATGTTGTTCA 8973
OY 63 GGCCTGAATTTTGTGCAGATAAATAAATAATCAATCA 101
DB 8974 GGGGTCACTCTTGTGTTAATAAATAATATATTTA 9012

RESULT 13
US-09-189-760-5/c
Sequence 5, Application US/09189760
Patent No. 6031078

GENERAL INFORMATION:
APPLICANT: Rhododoust, Mehran
TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: MNI-046CP2
CURRENT APPLICATION NUMBER: US/09/189,760
CURRENT FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 09/163,116
EARLIER FILING DATE: 1998-09-29
EARLIER APPLICATION NUMBER: 60/089,467
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: (PENDING)
EARLIER FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 5
LENGTH: 1529
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)..(749)
US-09-189-760-5

Query Match 24.2%; Score 25.6; DB 3; Length 1529;
Best Local Similarity 52.0%; Pred. No. 11;
Matches 52; Conservative 2; Mismatches 46; Indels 0; Gaps 0;

OY 4 ATTGACATTTTAAATAATATGTGAGTAGAGACTGCAGCTGTCAATAGCCTAG 63
DB 961 ATCTAATTTTAAAAAATGCTAGGTTGTTTCAGTTACTGCGACAAACAAAGCTTT 902
OY 64 GGCCTGAATTTTGTGCAGATAAATAAATAATCAATCA 103
DB 901 GGCACCTCTTTAGAGAAATTCACAAACAGATGCATC 862

RESULT 14
US-09-188-811-5/c
Sequence 5, Application US/09188811
Patent No. 6037148

GENERAL INFORMATION:
APPLICANT: Rhododoust, Mehran
TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: MNI-046CP
CURRENT APPLICATION NUMBER: US/09/188,811
CURRENT FILING DATE: 1998-11-09
EARLIER APPLICATION NUMBER: 09/163,116
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 5
LENGTH: 1529
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)..(749)

US-09-188-811-5

Query Match 24.2%; Score 25.6; DB 3; Length 1529;
Best Local Similarity 52.0%; Pred. No. 11;
Matches 52; Conservative 2; Mismatches 46; Indels 0; Gaps 0;

OY 4 ATTGACATTTTAAATAATATGTGAGTAGAGACTGCAGCTGTCAATAGCCTAG 63
DB 961 ATCTAATTTTAAAAAATGCTAGGTTGTTTCAGTTACTGCGACAAACAAAGCTTT 902
OY 64 GGCCTGAATTTTGTGCAGATAAATAAATAATCAATCA 103
DB 901 GGCACCTCTTTAGAGAAATTCACAAACAGATGCATC 862

RESULT 15
US-09-514-422-5/c
Sequence 5, Application US/09514422
Patent No. 6291193

GENERAL INFORMATION:
APPLICANT: Rhododoust, Mehran
TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: MNI-046CP2
CURRENT APPLICATION NUMBER: US/09/514,422
CURRENT FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: US/09/189,760
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: 09/163,116
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/089,467
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: (PENDING)
PRIOR FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 5
LENGTH: 1529
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)..(749)
US-09-514-422-5

Query Match 24.2%; Score 25.6; DB 4; Length 1529;
Best Local Similarity 52.0%; Pred. No. 11;
Matches 52; Conservative 2; Mismatches 46; Indels 0; Gaps 0;

OY 4 ATTGACATTTTAAATAATATGTGAGTAGAGACTGCAGCTGTCAATAGCCTAG 63
DB 961 ATCTAATTTTAAAAAATGCTAGGTTGTTTCAGTTACTGCGACAAACAAAGCTTT 902
OY 64 GGCCTGAATTTTGTGCAGATAAATAAATAATCAATCA 103
DB 901 GGCACCTCTTTAGAGAAATTCACAAACAGATGCATC 862

Search completed: April 3, 2002, 20:54:52
Job time: 41850 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:18:31 ; Search time 1321.64 Seconds
(without alignments)
68.760 Million cell updates/sec

Title: US-09-049-696-17

Perfect score: 106
Sequence: 1 GCGATTCACATTTTAAATAAT.....AAATAATCATTCATCCTT 106

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N.Geneseq.1101.*

1: /SIDSI/gcgdata/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	2825	22	AAH46124 Human C1CA1 cDNA.
2	106	100.0	2854	22	AAH34879 Human colon cancer
3	106	100.0	2854	22	AAH34879 Human secreted pro
4	106	100.0	2867	22	AAH33285 Human colon cancer
5	106	100.0	3109	22	AAH35019 Human colon cancer
6	106	100.0	3111	20	AAZ09840 Human membrane spa
7	106	100.0	3311	22	AAI29502 C902P determined c
8	103	97.2	255	16	AAI22483 Human gene signatu
9	76.8	72.5	878	18	AAI45884 Human colon specif
10	76.8	72.5	878	19	AAV16672 Polynucleotide seq
11	63	59.4	2745	20	AAH81927 Human ICACC-1 nucl

12	60	56.6	2742	22	AAH46102	Human C1CA1 coding
13	47.4	44.7	2843	20	AAH46120	Mouse Gob-5 cDNA.
14	47.4	44.7	2931	22	AAH81925	Murine ICACC-1 nuc
15	37.4	35.3	2739	22	AAH46101	Mouse Gob-5 coding
16	33.2	31.3	2520	15	AAO77884	Neural thread prot
17	33.2	31.3	2520	17	AAI27766	AD 16c human neu
18	32.4	30.6	7339	22	AAH59716	Human polynucleoti
19	29.6	27.9	7552	22	AAH04442	Human C1CA1 clone (
20	29	27.4	4321	22	AAO02314	N. tabacum arginin
21	29	27.4	3699	22	AAH55086	Nicotiana tabacum
22	28.6	27.0	129021	21	AAH22296	S. epidermidis gen
23	28.6	27.0	162025	22	AAH02339	BAC containing rep
24	28.6	27.0	162025	22	AAH02339	Human AKAP10 gene
25	28	26.4	2293	18	AAH89279	TPRP-F1 gene promo
26	28	26.4	6735	19	AAV52237	Streptococcus pneu
27	27.6	26.0	435	19	AAV57245	Clone #1 from muta
28	27.6	26.0	595	22	AAH34089	Human colon cancer
29	27.6	26.0	4436	9	AAH81716	Human growth hormo
30	27.6	26.0	134525	11	AAO04525	Total base sequenc
31	27.4	25.8	1123	19	AAV20912	Human DBI/ACBP -11
32	27.4	25.8	1337	22	AAI57930	Human polynucleoti
33	27.4	25.8	1560	22	AAH42601	Nucleotide sequenc
34	27.4	25.8	1574	22	AAH81572	Human endozepine-1
35	27.4	25.8	1887	21	AAH07584	Marigold epsilon-c
36	27.4	25.8	2243	22	AAH76407	F-box structure pr
37	27.4	25.8	6217	19	AAV65252	DNA encoding S. pn
38	27.4	25.5	221	22	AAI61519	Soybean 240017 reg
39	27	25.5	578	22	AAH13381	Human cDNA clone (
40	27	25.5	2674	21	AAZ56722	Human transmembran
41	27	25.5	4930	22	AAH18226	Human cDNA sequenc
42	27	25.5	13865	19	AAV40401	Human tissue facto
43	27	25.5	13865	20	AAZ32165	Human cholesterol
44	27	25.5	335913	22	AAI61371	Soybean 240017 reg
45	27	25.5	335913	22	AAI61372	Soybean 240017 reg

ALIGNMENTS

RESULT	1
ID	AAH46124 standard; cDNA; 2825 BP.
AC	AAH46124;
NC	AAH46124;
DE	11-SEP-2001 (first entry)
XX	Human C1CA1 cDNA, SEQ ID NO:26.
XX	Human C1CA1; goblet cell; mouse Gob-5 orthologue; drug screening;
KW	expression inhibition; antisense therapy; gene therapy;
KM	chronic obstructive pulmonary disease; bronchial asthma; antilasthmatic;
ES	
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	CD5
FT	Location/Qualifiers
FT	23..2767
FT	/*tag= a
FT	/product= "Human C1CA1"
FT	/trans_except= (pos:476..478, aa:lys)
XX	
XX	W0300138530-A1.
XX	
XX	31-MAY-2001.
XX	
XX	22-NOV-2000; 2000MO-JF08232.
XX	
XX	24-NOV-1999; 99JP-0333479.
XX	
XX	27-APR-2000; 2000JP-0127589.
XX	
XX	(TAKE) TAKEDA CHEM IND LTD.
XX	

PI Nakanishi A, Morita S;
XX WPI: 2001-355935/37.
DR P-PSDB: AAB73716.
XX
XX New antisense nucleotide, useful for treatment and prevention of
PT bronchial asthma and chronic obstructive pulmonary disease
XX
XX Example 5; Page 92-94; 104pp; Japanese.
XX
XX The invention relates to an antisense nucleotide targeted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the Gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
CC pulmonary disease. The present sequence represents human CLCA1 cDNA.
XX
XX Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;

Query Match 100.0%; Score 106; DB 22; Length 2825;
Best Local Similarity 100.0%; Pred. No. 1,7e-23;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGCATTCACATTTTAAATATGTCGAGAGTCGATAGAGAACGTCAGTCATAGCC 60
DB 2705 ggcattccattttaaataatctggaagtgagatagagactgcagctgcaataagcc 2764
OY 61 TAGGCTGCAATTTTGTGTCGATTAATAATATCATTCATCCCT 106
DB 2765 taggctgcaattttgtcagataataataatcattccctt 2810

RESULT 2
AAH34879
ID AAH34879 standard; cDNA; 2854 BP.
XX
XX AAH34879;
XX
XX 03-SEP-2001 (first entry)
XX
XX Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; chromosome 1; ss.
XX
XX Homo sapiens.
XX
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
XX
XX 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI: 2001-235357/24.
XX
XX P-PSDB: AAG73474.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1; Page 3462-3463; 9803pp; English.
XX

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match 100.0%; Score 106; DB 22; Length 2854;
Best Local Similarity 100.0%; Pred. No. 1,7e-23;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GGCATTCACATTTTAAATATGTCGAGAGTCGATAGAGAACGTCAGTCATAGCC 60
DB 2717 ggcattccattttaaataatctggaagtgagatagagactgcagctgcaataagcc 2776
OY 61 TAGGCTGCAATTTTGTGTCGATTAATAATATCATTCATCCCT 106
DB 2777 taggctgcaattttgtcagataataataatcattccctt 2822

RESULT 3
AAF81787
ID AAF81787 standard; cDNA; 2854 BP.
XX
XX AAF81787;
XX
XX 12-JUN-2001 (first entry)
XX
XX Human secreted protein gene 1 SEQ ID NO:11.
XX
XX Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
XX dermatological; immunosuppressive; anti-inflammatory; anti-HIV;
XX immunostimulant; cytostatic; cardiac; vascular; anti-angiogenic;
XX opthalmochemical; neuroprotectant; nootropic; anticonvulsant; vaccine;
XX anti-Alzheimer; antiparkinsonian; antimicrobial; vulnery; gene therapy;
XX immune disorder; hyperproliferative disorder; cardiovascular disease;
XX cancer; angiogenic disorder; neurological disorder; infectious disease;
XX wound healing; regeneration; chemotaxis; chromosome 1; ss.
XX
XX Homo sapiens.
XX
XX WO200112775-A2.
XX
XX 22-FEB-2001.
XX
XX 16-AUG-2000; 2000WO-US22325.
XX
XX 17-AUG-1999; 99US-0149182.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Resen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;
XX
XX Birse CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;
XX
XX WPI: 2001-147550/15.
XX
XX P-PSDB: AAB74733.
XX

PT Nucleic acids encoding 25 human secreted polypeptides, useful for
preventing, diagnosing and/or treating e.g. cancers, Parkinson's
disease and diabetic retinopathy.

PS Claim 1, Page 441; 485pp; English.

CC AAF81787 to AAF81817 encode the human secreted proteins given in AAF81773
CC to AAF81772. Human secreted proteins can have activities based on the
CC tissues and cells they are expressed in. Example of activities include:
CC immunomodulatory; antileukemic; dermatological; immunosuppressive;
CC antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;
CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
CC anxiolytic; antialzheimer's; antiparkinsonian; antimicrobial; and
CC vaccine. Human secreted proteins can be used in gene therapy and
CC vullnary. Human secreted protein nucleotide sequences (NMI) and proteins
CC (PPI) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. For example, NMI
CC and PPI may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of proteins by expressing inactive proteins or
CC to supplement the patient's own production of polypeptides. Disorders that
CC may be prevented, diagnosed and/or treated include immune disorders,
CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
CC angiogenic disorders, neurological disorders, infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to
CC AAF81786 and AAF81732 represent sequences used in the exemplification of
CC the present invention.

SO Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match 100.0%; Score 106; DB 22; Length 2854;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTCACATTTTAAATATGTCGAGTGTGGAAGTGTGCAATGACC 60
DB 2717 ggcattcacattttaaataatgtagagtgataggaactgcagctcatalagcc 2776
OY 61 TAGGGCTGAATTTTGTGTCGATTAATTAATTAATCATTCATCCCTT 106
DB 2777 tagggctgaattttgttcgataataataataatcattccctt 2822

RESULT 4

AAH33285
ID AAH33285 standard; cDNA; 2867 BP.

AC AAH33285;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:341.

KM Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 1; ss.

OS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI: 2001-235357/24.

DR P-PSDB: AAG73854.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -

PS Claim 1, Page 2452-2453; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing P.
CC Inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC present invention.

CC N/A: Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

SO Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;

Query Match 100.0%; Score 106; DB 22; Length 2867;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTCACATTTTAAATATGTCGAGTGTGGAAGTGTGCAATGACC 60
DB 2721 ggcattcacattttaaataatgtagagtgataggaactgcagctcatalagcc 2780
OY 61 TAGGGCTGAATTTTGTGTCGATTAATTAATTAATCATTCATCCCTT 106
DB 2781 tagggctgaattttgttcgataataataataatcattccctt 2826

RESULT 5

AAH35019
ID AAH35019 standard; cDNA; 3109 BP.

AC AAH35019;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:2101.

KM Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.

OS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI: 2001-235357/24.

XX P-PSDB: AAG75614.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers.
 XX
 XX Claim 1; Page 3587-3588; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match 100.0%; Score 106; DB 22; Length 3109;
 Best Local Similarity 100.0%; Pred. No. 1.7e-23;
 Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTCACATTTTAAATAATATGCGAAGTGGAGTGAAGAGAGAGCTGTCATATAGCC 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2570 ggcattcacattttaaaataatcattggaagtggaaggaacgcgctgcaatagcc 2629
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

OY 61 TAGGCGTGAATTTTGTGACATTAATAATAATCATTCATTCCTT 106
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2630 tagggctgaattttgtcagataataataataatcattcattcctt 2675
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 6
 AAZ09840
 ID AAZ09840 standard; cDNA; 3111 BP.
 XX
 AC AAZ09840;
 XX
 DT 26-NOV-1999 (first entry)
 XX
 DE Human membrane spanning protein MSP-5 cDNA fragment 2.
 XX
 DE Human membrane spanning protein; MSP: human; treatment; diagnosis; prevention;
 KM neoplastic disorder; immunological disorder; reproductive disorder;
 KM MSP-5; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO9946380-A2.
 XX
 PD 16-SEP-1999.
 XX
 PF 09-MAR-1999; 99WO-US05073.
 XX
 PR 13-MAR-1998; 98US-0039064.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Tang YT, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;
 PI Guegler KJ, Kaser MR, Baughn MR, Shah P;
 XX WPI: 1999-551409/46.
 DR P-PSDB; AAZ33298.

XX New human membrane spanning proteins used to, e.g. prevent and treat
 PT neoplastic disorders -
 XX
 XX Example 1; Page 80-81; 81pp; English.
 PS
 CC This invention describes novel human membrane spanning proteins (MSPs),
 CC and the polynucleotides encoding them. The products of the invention are
 CC used to diagnose, prevent and treat neoplastic, immunological and
 CC reproductive disorders. This sequence encodes a human membrane spanning
 CC protein MSP-5 fragment.
 XX
 SQ Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match 100.0%; Score 106; DB 20; Length 3111;
 Best Local Similarity 100.0%; Pred. No. 1.7e-23;
 Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTCACATTTTAAATAATATGCGAAGTGGAGTGAAGAGAGAGCTGTCATATAGCC 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2716 ggcattcacattttaaaataatcattggaagtggaaggaacgcgctgcaatagcc 2775
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

OY 61 TAGGCGTGAATTTTGTGACATTAATAATAATCATTCATTCCTT 106
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2776 tagggctgaattttgtcagataataataataatcattcattcctt 2821
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 7
 AAI29502
 ID AAI29502 standard; cDNA; 3311 BP.
 XX
 AC AAI29502;
 XX
 DE 12-OCT-2001 (first entry)
 XX
 DE C902P determined cDNA sequence.
 XX
 KM Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 KM immunogenic; gene therapy; vaccine; colonic cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WC200149716-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 29-DEC-2000; 2000WO-US35596.
 XX
 PR 30-DEC-1999; 99US-0476286.
 XX
 PR 10-JAN-2000; 2000US-0480321.
 XX
 PR 15-FEB-2000; 2000US-0504629.
 XX
 PR 06-MAR-2000; 2000US-0519444.
 XX
 PR 19-MAY-2000; 2000US-0575251.
 XX
 PR 29-JUN-2000; 2000US-0609448.
 XX
 PR 28-AUG-2000; 2000US-0649811.
 XX
 PA (CORI-) CORINA CORP.
 XX
 PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
 PI King GE, Wang T, Jiang Y;
 XX WPI: 2001-441847/47.
 XX
 CC Colon tumor associated proteins and nucleic acids useful for the
 CC prevention, diagnosis and treatment of colonic cancer.
 XX
 CC Claim 2; Page 425-426; 472pp; English.
 XX
 CC The present invention describes colon tumour associated proteins (I) and
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases

xx 13 cDNA clones (AA145880-92), most of them partial clones, correspond
 CC to human colon specific genes, designated CSG1, CSG2, etc., that
 CC are primarily expressed in tissues derived from the colon. CSG7
 CC and CSG10 show reduced expression in colon cancer cells as compared
 CC to that in normal cells; the remaining genes are overexpressed in
 CC colon cancer. The partial cDNA sequences can be used to isolate
 CC full-length clones and genomic clones including the complete gene.
 CC CSG nucleic acids can be used to produce CSG polypeptides (see also
 CC AAW06545-53) in transformed host cells, as probes to detect disorders
 CC of the colon, partic. colon cancer and colon cancer metastasis, and
 CC in gene therapy.

SO Sequence 878 BP; 257 A; 179 C; 188 G; 241 T; 13 other;

Query Match 72.5%; Score 76.8; DB 18; Length 878;
 Best Local Similarity 91.0%; Pred. No. 1.3e-14;
 Matches 81; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 GGCATTCACTTTTAAATATGATGAGATGAGAGAACTGCTCATACCC 60
 |||||||
 DB 717 ggcattcactttaaaatattatggaagtggtaggaagactgcagttgcaatagnc 776
 |||||||

OY 61 TAGGCTCAATTTTGTCTAGATAATAA 89
 |||||
 DB 777 taggggtgaattttgtcgtgtaataa 805
 |||||

RESULT 10

AAV16672
 ID AAV16672 standard; cDNA; 878 BP.

AC AAV16672;

DT 22-JUN-1998 (first entry)

XX Polynucleotide sequence of a colon-specific gene.

XX Colon-specific gene; probe; detection; expression; human;
 KW diagnostic assay; colon cancer; antibody; screening; ss.

OS Homo sapiens.

EH Key Location/Qualifiers
 FT CDS 2..685
 FT /*tag= a
 FT /note= "no stop codon given"

PN US5733748-A.

PD 31-MAR-1998.

PE 06-JUN-1995; 95US-0469667.

PR 06-JUN-1995; 95US-0469667.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen C, Yu G;

DR WPI: 1998-229823/20.

DR P-PSDB; AAW46879.

XX Colon-specific nucleic acids - useful as probes for detecting colon
 PT cancer micrometastases

PS Claim 15; Fig 5A-B; 51pp; English.

XX AAV16668-81 represent polynucleotide sequences of partial or full-length
 CC cDNA clones of colon-specific genes. The polynucleotides can be used
 CC as probes to detect expression of the corresponding human genes, e.g. in
 CC diagnostic assays for detecting micrometastases of colon cancer.

CC Recombinant cells containing the polynucleotides can be used to
 CC produce the polypeptides, in order that antibodies can be raised and
 CC used in further screening or diagnostics.

SO Sequence 878 BP; 257 A; 179 C; 188 G; 241 T; 13 other;

Query Match 72.5%; Score 76.8; DB 19; Length 878;
 Best Local Similarity 91.0%; Pred. No. 1.3e-14;
 Matches 81; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 GGCATTCACTTTTAAATATGATGAGATGAGAGAACTGCTCATACCC 60
 |||||||
 DB 717 ggcattcactttaaaatattatggaagtggtaggaagactgcagttgcaatagnc 776
 |||||||

OY 61 TAGGCTCAATTTTGTCTAGATAATAA 89
 |||||
 DB 777 taggggtgaattttgtcgtgtaataa 805
 |||||

RESULT 11

AAV81927
 ID AAV81927 standard; cDNA; 2745 BP.

AC AAV81927;

DT 13-JUN-2001 (first entry)

XX Human ICACC-1 nucleotide sequence.

XX ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
 KW interleukin 9 induced calcium activated chloride channel; IL-9;
 KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
 KW anti-inflammatory; immunomodulatory; cystic fibrosis;
 KW inflammatory bowel disease; autoimmune disease; ss.

OS Homo sapiens.

EH Key Location/Qualifiers
 FT CDS 1..2745
 FT /*tag= a
 FT /product= "ICACC-1"
 FT /note= "IL-9 induced calcium activated chloride channel"

PN WC8944620-A1.

PD 10-SEP-1999.

PE 03-MAR-1999; 99WO-US04703.

PR 03-MAR-1998; 98US-0076815.

PA (MAGA-) MAGATININ PHARM INC.

PI Holroyd KJ, Levitt RC, Maloy WL, Louahed J, McLane M;
 PI Nicolaidis NC, Zhou Y, Dong Q;

DR WPI: 1999-550979/46.

DR P-PSDB; AAV74824.

XX New nucleic acid encoding calcium activated chloride channel, used to
 PT identify, e.g. specific modulators for treating atopic allergy -
 XX Claim 1; Fig 4B; 75pp; English.

XX The present sequence encodes the human interleukin 9 (IL-9) induced
 CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
 CC have anti-allergic, anti-asthmatic, anti-inflammatory and
 CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
 CC used to alleviate asthma (or more generally atopic allergy), while those
 CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
 CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
 CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma

CC (increased levels) or IBD (reduced levels), also for monitoring
 CC treatment of these conditions. The ICAC proteins can be used:
 CC (i) to raise specific antibodies (Ab), useful: (a) as immunoassay
 CC reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
 CC (A) to disrupt bonding between ICAC and its ligands; and (iii) to
 CC identify modulators and binding partners. ICAC polynucleotides can be
 CC used to generate transgenic animals or recombinant cells, used to screen
 CC for antagonists, also as a source of therapeutic antisense agents or
 CC diagnostic probes (for quantifying mRNA expression, e.g. for
 CC identification of modulators).

XX Sequence 2745 BP: 833 A: 616 C: 625 G: 671 T: 0 other:

Query Match 59.4%; Score 63; DB 20; Length 2745;
 Best Local Similarity 100.0%; Pred. No. 2.9e-10;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTACATTTTAAATTTATGTGAGTAGAGAGACTGACGTCAATAGCC 60
 |||||||
 DB 2683 ggcattcacattttaaataattatgtggaagtgtagtaggaactgcagctcacaatagcc 2742
 OY 61 TAG 63
 |||
 DB 2743 tag 2745

RESULT 12

AAH46102 standard; DNA; 2742 BP.

AAH46102;

11-SEP-2001 (first entry)

Human CLCA1 coding sequence. SEQ ID NO:4.

Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;

expression inhibition; antisense therapy; gene therapy;
 chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
 ds.

Homo sapiens.

Key Location/Qualifiers

CDS 1..2742

/tag= a

/product= "Human CLCA1"

/note= "No stop codon given in the specification"

MO200138530-A1.

31-MAY-2001.

22-NOV-2000; 2000MO-JP08232.

24-NOV-1999; 99JP-0333479.

27-APR-2000; 2000JP-0127589.

(TAKE) TAKEDA CHEM IND LTD.

Nakanishi A, Morita S;

WPI: 2001-355935/37.

P-PSDB: AAB73716.

New antisense nucleotide, useful for treatment and prevention of
 bronchial asthma and chronic obstructive pulmonary disease -

Claim 3; Page 82-84; 104pp; Japanese.

The invention relates to an antisense nucleotide targeted to the mouse

CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
 CC relates to an antibody specific for the Gob-5 protein, medical and
 CC diagnostic compositions containing the antisense nucleotide or the
 CC antibody, and methods and kits for screening for compounds which inhibit
 CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 CC The antisense oligonucleotides and antibody are therefore useful for the
 CC treatment and prevention of bronchial asthma and chronic obstructive
 CC pulmonary disease. The present sequence represents the human CLCA1
 CC gene coding sequence.

XX Sequence 2742 BP: 833 A: 616 C: 623 G: 670 T: 0 other:

Query Match 56.6%; Score 60; DB 22; Length 2742;
 Best Local Similarity 100.0%; Pred. No. 2.4e-09;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTACATTTTAAATTTATGTGAGTAGAGAGACTGACGTCAATAGCC 60
 |||||||
 DB 2683 ggcattcacattttaaataattatgtggaagtgtagtaggaactgcagctcacaatagcc 2742

RESULT 13

AAH46120 standard; cDNA; 2843 BP.

AAH46120;

11-SEP-2001 (first entry)

Mouse Gob-5 cDNA, SEQ ID NO:22.

Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
 expression inhibition; antisense therapy; gene therapy; bronchial asthma;
 chronic obstructive pulmonary disease; antiasthmatic; ss.

Mus sp.

Key Location/Qualifiers

CDS 15..2756

/tag= a

/product= "Mouse Gob-5"

MO200138530-A1.

31-MAY-2001.

22-NOV-2000; 2000MO-JP08232.

24-NOV-1999; 99JP-0333479.

27-APR-2000; 2000JP-0127589.

(TAKE) TAKEDA CHEM IND LTD.

Nakanishi A, Morita S;

WPI: 2001-355935/37.

P-PSDB: AAB73715.

New antisense nucleotide, useful for treatment and prevention of
 bronchial asthma and chronic obstructive pulmonary disease -

Example 1; Page 89-91; 104pp; Japanese.

The invention relates to an antisense nucleotide targeted to the mouse
 Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
 CC relates to an antibody specific for the Gob-5 protein, medical and
 CC diagnostic compositions containing the antisense nucleotide or the
 CC antibody, and methods and kits for screening for compounds which inhibit
 CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.

The invention relates to an antisense nucleotide targeted to the mouse
 Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
 CC relates to an antibody specific for the Gob-5 protein, medical and
 CC diagnostic compositions containing the antisense nucleotide or the
 CC antibody, and methods and kits for screening for compounds which inhibit
 CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.

CC relates to an antibody specific for the Gob-5 protein, medical and
 CC diagnostic compositions containing the antisense nucleotide or the
 CC antibody, and methods and kits for screening for compounds which inhibit
 CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 CC The antisense oligonucleotides and antibody are therefore useful for the
 CC treatment and prevention of bronchial asthma and chronic obstructive
 CC pulmonary disease. The present sequence represents the mouse Gob-5
 CC gene coding sequence.

XX
 SQ Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other;

Query Match 35.38; Score 37.4; DB 22; Length 2739;
 Best Local Similarity 74.68; Pred. No. 0.022;
 Matches 47; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 GGCATTGCACATTTTAAATAATTATGTGAGTGCATAGCAGACACGCTCTCAATAGCC 60
 Db 2674 ggcattgcacgtgtgaagataatgtgagctaggggaaatcagtgacactagt 2733

QY 61 TAG 63
 Db 2734 ttg 2736

Search completed: April 4, 2002, 03:18:34
 Job time: 60266 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:07:59 ; Search time 22700.8 Seconds
(without alignments)
77.033 Million cell updates/sec

Title: US-09-049-696-17
Perfect score: 106
Sequence: 1 GGCATTGCATTTTAAAT.....AAATTAATCATTCATCCTT 106

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8 SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	106	100.0	2022	9	AK024970	AK024970 Homo sapi
2	106	100.0	2826	9	AF127036	AF127036 Homo sapi
3	106	100.0	3311	6	AX193489	AX193489 Sequence
4	106	100.0	3311	9	AF039400	AF039400 Homo sapi
5	106	100.0	35278	9	AF039401	AF039401 Homo sapi
6	106	100.0	113764	9	HS1651E10	AL122002 Homo sapi
7	106	100.0	164891	2	AL358950	AL358950 Homo sapi
8	72.5	878	6	195746	AF095584	AF095584 Sus scrofa
9	71.4	3079	4	AF095584	AB017156	AB017156 Mus musculus
10	47.4	2937	10	AB017156	BSCHEVE	229584 B.subtilis
11	36.6	34.5	3215	1	BSU05345	005345 Bacillus subtilis
12	36.6	34.5	3215	1	BS16829KB	AL222587 Bacillus subtilis
13	36.6	34.5	29324	1	BSUB0008	299111 Bacillus subtilis
14	36.6	34.5	208230	1	BSUB0008	299286 Caenorhabditis elegans
15	34	32.1	12984	1	AE001122	AE001122 Borrelia burgdorferi
16	34	32.1	26821	3	CEY7A9C	299286 Caenorhabditis elegans
17	34	32.1	110000	2	CEY113B8_1	Continuation (2 of 272517 Caenorhabditis elegans)
18	33.6	31.7	30567	3	CEY113B8_1	Continuation (2 of 272517 Caenorhabditis elegans)
19	33.4	31.5	83538	2	AC018338	AC018338 Drosophila melanogaster
20	33.4	31.5	249805	3	AE003515	AE003515 Drosophila melanogaster
21	33.2	31.3	2520	6	AR051480	AR051480 Sequence
22	33.2	31.3	2520	6	AR072620	AR072620 Sequence
23	33.2	31.3	2520	6	AR073165	AR073165 Sequence
24	33	31.1	5460	8	AB032938	AB032938 Cucumis sativus
25	32.8	30.9	171631	2	AC069291	AC069291 Homo sapi
26	32.8	30.9	171788	9	AC073089	AC073089 Homo sapi
27	32.8	30.9	189993	2	AC079920	AC079920 Homo sapi
28	32.4	30.6	2508	1	STRSCRA	M22711 Streptococcus thermophilus
29	32.2	30.4	60271	2	AC090212	AC090212 Homo sapi
30	32.2	30.4	164064	9	AL138879	AL138879 Human DNA
31	32.2	30.4	193530	2	AP003455	AP003455 Oryza sativa
32	32.2	30.4	270795	2	AC015535	AC015535 Mus musculus
33	32	30.2	154478	8	AP000423	AP000423 Arabidopsis thaliana
34	32	30.2	175789	2	AC090425	AC090425 Homo sapi
35	32	30.2	202959	2	AC091738	AC091738 Homo sapi
36	31.8	30.0	103916	9	AC004225	AC004225 Homo sapi
37	31.8	30.0	106556	9	HS0512E2	AL121787 Human DNA
38	31.8	30.0	122308	2	AC016553	AC016553 Homo sapi
39	31.8	30.0	123241	2	AC022090	AC022090 Homo sapi
40	31.6	29.8	175916	2	AC007747	AC007747 Homo sapi
41	31.6	29.8	110000	2	AC087899_2	Continuation (3 of 195746 Sequence 8)
42	31.6	29.8	160831	2	AC025015	AC025015 Homo sapi
43	31.6	29.8	161130	2	AC021050	AC021050 Homo sapi
44	31.6	29.8	163232	2	AC084017	AC084017 Homo sapi
45	31.6	29.8	174555	2	AC023872	AC023872 Homo sapi

ALIGNMENTS

RESULT	1	2
AK024970	AK024970	AK024970
LOCUS	Homo sapiens cDNA: FLJ21317 fls, clone COL02275, highly similar to	PRI 29-SEP-2000
DEFINITION	AF127036 Homo sapiens calcicium-activated chloride channel protein 1 (CACCL1) mRNA.	
ACCESSION	AK024970	GI:10437397
VERSION	AK024970.1	GI:10437397
KEYWORDS	oligo capping; fls (full insert sequence).	
SOURCE	Homo sapiens colon cDNA to mRNA, clone_11b:COL clone:COL02275.	
ORGANISM	Homo sapiens	
REFERENCE	1 (sites)	
AUTHORS	Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.	
TITLE	NEO human cDNA sequencing project	
JOURNAL	Unpublished (2000)	
REFERENCE	2 (bases 1 to 2022)	

AUTHORS Sugano, S., Suzuki, Y., Ota, T., Ohayashi, M., Nishi, T., Isogai, T., Shihahara, T., Tanaka, T. and Nakamura, Y.

JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/Genbank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: chnail@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

COMMENT

NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' and 3' end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source 1..2022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL02275"
/clone_11b="COL"
/tissue="colon"
/note="Cloning vector pME18SFL3"
1..2022
/note="highly similar to AFI27036 Homo sapiens calcium-activated chloride channel protein 1 (CACCL) mRNA"

BASE COUNT 612 a 472 c 453 g 485 t

ORIGIN

Query Match 100.0%; Score 106; DB 9; Length 2022;
Best Local Similarity 100.0%; Pred. No. 1.2e-20;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTCACTTTTAAATATGAGAGTGAAGAGAACTGACGCTGCAATAGCC 60
|||||
Db 1896 GGCATTCACTTTTAAATATGAGAGTGAAGAGAACTGACGCTGCAATAGCC 1955
|||||

OY 61 TAGGGCTGAATTTTGTGAGATAAATAATATCATTCATCCTT 106
|||||
Db 1956 TAGGGCTGAATTTTGTGAGATAAATAATATCATTCATCCTT 2001
|||||

RESULT

2

AF127036 2826 bp mRNA PRI 10-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 1 (CACCL)
DEFINITION mRNA, complete cds.
ACCESSION AF127036
VERSION AF127036.1 GI:4585468
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 2826)

Agnel, M., Vernat, T. and Culouscou, J. M.

TITLE

JOURNAL

FEBS Lett. 455 (3), 295-301 (1999)

MEDLINE

PUBMED

2 (bases 1 to 2826)

Agnel, M. and Culouscou, J. M.

TITLE

JOURNAL

Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrières, Rueil-Malmaison 92500, France

location/Qualifiers

1..2826

/organism="Homo sapiens"

/db_xref="taxon:9606"

gene
/tissue="small intestine; colon"
1..2826
/gene="CACCL"

CDS

5..2749

/gene="CACCL"

/note="bovine epithelial chloride channel homolog"

/product="calcium-activated chloride channel protein 1"

/protein_id="A025487.1"

/db_xref="GI:4585469"

/translation="MGRFSSVYILHLLEGALSLGIQNNNGEIVAVIDPVP
EDTLTQIKIDMTQASLVLEFYTGKRFKNAIIPETWTKADYRPKLETKNA
DVAIASTPGNDPEYEQNGCEGRIHLIPDI IAKKLAIEGPGRAVWEVAH
LWGVDEYNNDEKFTLSNGRIQAVCSAGITGVYKCGGCTGKRCITNKVGL
YKGCERVLQRTOKASIMFAQVDSIYFCEQHNKAPKONKCNLSTWEL
RDSFDRKTPMTTOPNPPKSEILOIGRIVCLVILSGSMATGNTLNAGOLEL
LQTVELSGVGMATFDSAAHVQSELIQINSGRDPLARIPAAAGCTSGISLPSA
FTVIRKKYPTDGSSEIVLTDGENTLSGCNEVKSQGAIIHYVALGPSAQLLELSK
MTGGLQVYASDOVQNNGLIDAFGLSSNGVAVSRIQLESKLTILQNSOMNGTIV
DSYVGDITLFLITWTPQIILMDPSGKQGGVVDKTKNAVYQIIP IAKGTWKY
SLQASSQTLTIVTSRASNAITLPITVTSKNTDTSKFPSPLVVANIHQASPIILRA
SYVALIESVNGKVTLEILNDGAGADTKDGYSHRYFTYDNGRYSVKRALGAVN
AARRVYPOOSGALYIPGTEENDEIOMNPPRIKNDPOHKOVCSPRSGGSPVAS
DVPNAPLDLPPEQITDLKAEIHGSLINUTATAGDDYDHTAKYIIRISTILD
LRDFNLSLVNVTALIPKEANSSEVLEFPEVITEENGTDLFIATQADKDLKSEI
SNIAVSLFIPQTPPETPSPDETSAPCPNHNINSTIPGIHLIKMKWIGELQSLIA

BASE COUNT 875 a 623 c 632 g 696 t

ORIGIN

Query Match 100.0%; Score 106; DB 9; Length 2826;
Best Local Similarity 100.0%; Pred. No. 1.2e-20;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCATTCACTTTTAAATATGAGAGTGAAGAGAACTGACGCTGCAATAGCC 60
|||||
Db 2637 GGCATTCACTTTTAAATATGAGAGTGAAGAGAACTGACGCTGCAATAGCC 2746
|||||

OY 61 TAGGGCTGAATTTTGTGAGATAAATAATATCATTCATCCTT 106
|||||
Db 2747 TAGGGCTGAATTTTGTGAGATAAATAATATCATTCATCCTT 2792
|||||

RESULT

3

AX193489 3311 bp DNA PAT 15-AUG-2001
LOCUS Sequence 1056 from Patent WO0149716.
ACCESSION AX193489
VERSION AX193489.1 GI:15211440
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 3311)

Xu, J., Lodes, M. J., Secrist, H., Benson, D. R., Meagher, M. J., Stolk, J. A., King, G. E., Wang, T. and Jiang, Y.

TITLE

JOURNAL

Compounds for immunotherapy and diagnosis of colon cancer and Patient: WO 0149716-A 1056 12-JUL-2001;

CORIXA CORPORATION (US)

location/Qualifiers

1..3311

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT

1028 a 692 c 742 g 849 t

ORIGIN

Query Match 100.0%; Score 106; DB 6; Length 3311;
Best Local Similarity 100.0%; Pred. No. 1.3e-20;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCATTACATTTTAAATTAATGTCGACGTGATGAGAGACTGCAGTGTCAATGCC 60
|||||
DB 3034 GGCATTACATTTTAAATTAATGTCGACGTGATGAGAGACTGCAGTGTCAATGCC 3093
QY 61 TAGGGCTGAATTTTGTGCAGATTAATAATTAATCATTCATCCTT 106
|||||
DB 3094 TAGGGCTGAATTTTGTGCAGATTAATAATTAATCATTCATCCTT 3139
RESULT 4
AF039400 3311 bp mRNA PRI 14-DEC-1998
LOCUS AF039400
DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA,
complete cds.
ACCESSION AF039400
VERSION AF039400.1 GI:4009457
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Gruber,A.D., Eblie,R.C., Ji,H.L., Schreur,K.D., Fuller,C.M. and
Pauli,B.U.
TITLE Genomic cloning, molecular characterization, and functional
analysis of human CLCA1, the first human member of the family of
Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 3311)
AUTHORS Gruber,A.D., Eblie,R. and Pauli,B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
Cornell University, Ithaca, NY 14853, USA
FEATURES
source
1..3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1p22-p31"
/tissue_type="small intestine"
1..3311
/gene="hCLCA1"
352..3096
/gene="hCLCA1"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPKSVFIIILHLEGLASLSLIQLNNNGEIVADPNV
EDETLLOIKMVTQASLTFRATGRREFKRVAYILIPETWTKADYVRPKLETYNK
DYLVAESTPPGNDPEPTTEOMGCKEGERIHLPDIACKLAETGPQKRAVEHNAH
LWGVFDEYNDEKFIYLSNGRIQAVKCSAGITGVNKKCGSCCTKCTKRVKGL
YKGECEFYVLOSOTEKASIMFAOHVDSIYEFTGONHNKAPNKKONKLNSTWYI
RSEDEPKTTPMTQPPNPFTSLQIGRIYGLVDKSGSMATGNLNINLNGAGOLF
LOTVELGSWVGWTFDSAHVQSELIQINSGSDRLAKRLPAASGGNSGSGSLASA
FTVIRKKYPTGSEIYLLTDGNTLSCGFENYKOGAIIHVTALPSAAOELBELSK
MTGGLQTVASDVOYNNGLDANGLASSGGAVSORSIQLESGLILQNSOMNGTVIV
DSTVGKDLLELTWTQPPQILLMPDSGGKGFVVDKTKMAVLDIPGIAVGTWKY
SLAASSQTLTFTVTSRASNATLPPIVTSKTKNDISKPSPIVYVNIQOGASPIRA
SVTALIESVNGKTVLELDNGAGADATGDCVSYRFTYDNGRYSKVRALGGVA
AARRVYIPQOSGALYIPGWIEDEIOMNPPEINDDYOHKOVCSRPSGSEFVAS
DVNAPIPDLPFGOITDKLEIHGSLINLWTVAGDDYDHTAKYIIRISTSLD
LADKRNESLQVNTTALITREANSSEVFLKPEKITFENGSTDLFTAIQAVDKVLDKEI
SNIAVSLFIPQTPPEPSPDSTSPCNHINSTIPGIIHLIKMKWIGELQLSIA

Query Match

100.0%; Score 106; DB 9; Length 3311;

Best Local Similarity 100.0%; Pred. No. 1,3e-20;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCATTACATTTTAAATTAATGTCGACGTGATGAGAGACTGCAGTGTCAATGCC 60
|||||
DB 3034 GGCATTACATTTTAAATTAATGTCGACGTGATGAGAGACTGCAGTGTCAATGCC 3093
QY 61 TAGGGCTGAATTTTGTGCAGATTAATAATTAATCATTCATCCTT 106
|||||
DB 3094 TAGGGCTGAATTTTGTGCAGATTAATAATTAATCATTCATCCTT 3139
RESULT 5
AF039401 35278 bp DNA PRI 14-DEC-1998
LOCUS AF039401
DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene,
complete cds.
ACCESSION AF039401
VERSION AF039401.1 GI:4009459
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Gruber,A.D., Eblie,R.C., Ji,H.L., Schreur,K.D., Fuller,C.M. and
Pauli,B.U.
TITLE Genomic cloning, molecular characterization, and functional
analysis of human CLCA1, the first human member of the family of
Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 35278)
AUTHORS Gruber,A.D., Eblie,R. and Pauli,B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Department of Pathology, College of
Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES
source
1..35278
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
1..1617
/gene="hCLCA1"
1..33522
/gene="hCLCA1"
767..1265
/rpl_family="L1 transposable element"
/rpl_type-dispersed
1587..1596
/gene="hCLCA1"
/note="putative"
1618..1708
/gene="hCLCA1"
/number=1
join(1618..1708,1962..2383,6667..6807,6979..7126,
9696,19801,15454,18809,19783..20010,22253..22427,
24522..24628,26644..26859,27447..27708,28766,
28936,31840..32079,32919..33522)
/gene="hCLCA1"
/product="calcium-dependent chloride channel-1"
1962..2383
/gene="hCLCA1"
/number=2
join(2222..2383,6667..6807,6979..7126,9696,
15454..15631,18591..18809,19783..20010,22253..
24522..24628,26644..26859,27447..27708,28766,
28936,31840..32079,32919..33510)
/gene="hCLCA1"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95429.1"


```

/db_xref="GI:4009460"
/translation="MGPFKSSVFILILHLEGLASNSLTOLNNNGECIVADPNP
EDFTLQIQIDMVTQASLYLEATGKRFRKMAILLPEWTKRADYVRKLEFYKA
DVLVAESTPGNDPEYTEMONGNCEKGERIHLIPDFIAGKRLAEGPOGALEMAH
LRMGVPEYNNDEKFLYNGRIQAVRSGEITGTNNVYKCGGCGYKRIKFNKVTGL
YKCGCEVLOSROTERKASIMPAOHVDSIVCECTEONKNEAPNOKCKMLRSTWEI
RDSDEFKTTPMTOPNPPEPSILQIGORIVCIAYLDSGSMATGNRLNINOGDFTL
LQVELSGWGVTFDSNAHVOSELIOINGSDPDLAKLPAASGCTSCGIBSA
FTVIRKRPDGESEIVLLTDEGNTISGCENEVKQSAIIHVALGSAOEBELSK
MTGLQTYASDOVONGLIDAFGLSSGNVAVSQRIOLESKGLTTONSOMNCTIV
DSYVCKDLELITWTQPOLILMDPSGOGKGVVKNKRMAYLIQIGAFKNTVL
SLQASQDILITVSRASNAIPLPIVTSKNTKDSKEPSPLVYANIRGASPIILA
SVTALISVNCVTLELLNGAGADATKDGVSRYETTYDNGRYSVYRABALGN
AARRVTPQSGALITPGWINDETIQNNRPPEINKDVOHKOCESRISGSEFVAS
DVNPAPIDLPFPPQITDIAEIHGSLINLTWAPDDYDHGAHKYITIRISTIID
LRDKFNSLOVNTALIPKEANSEVEFLFRENTEFNGDFTLAIQAVDKVLSKSI
SNIRVSLFIPQTPPEPSPDETSACPNIHINSTIPGIHLIKMKWIGELQSLIA
"
exon      6667..6807
           /gene="hclcal1"
           /number=3
exon      6979..7126
           /gene="hclcal1"
           /number=4
exon      9696..9801
           /gene="hclcal1"
           /number=5
exon      15454..15631
           /gene="hclcal1"
           /number=6
exon      18591..18809
           /gene="hclcal1"
           /number=7
exon      19783..20010
           /gene="hclcal1"
           /number=8
exon      22253..22427
           /gene="hclcal1"
           /number=9
exon      24522..24628
           /gene="hclcal1"
           /number=10
exon      26644..26859
           /gene="hclcal1"
           /number=11
exon      27447..27708
           /gene="hclcal1"
           /number=12
exon      28766..28936
           /gene="hclcal1"
           /number=13
exon      31840..32079
           /gene="hclcal1"
           /number=14
exon      32919..33522
           /gene="hclcal1"
           /number=15
BASE COUNT 11245 a 6870 c 7230 g 9933 t
ORIGIN
Query Match      100.0%; Score 106; DB 9; Length 35278;
Best local Similarity 100.0%; Pred. No. 1.5e-20;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GGCATTACATTAAATAATATGATGAGAGTGGAGAACTGACGCTCAATAGCC 60
        |||||||
DB      33248 GGCATTACATTAAATAATATGATGAGAGTGGAGAACTGACGCTCAATAGCC 33307
        |||||||
QY      61 TAGGGCTGAATTTTGTGAGATAAATAATATCAATCATTCCTT 106
        |||||||
DB      33308 TAGGGCTGAATTTTGTGAGATAAATAATATCAATCATTCCTT 33353

```

```

RESULT: 6
HSJ651E10
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: hunquerry@sanger.ac.uk
On Jun 4, 2000 this sequence version replaced gi:8247030.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; SW, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP4-651E10 is from the library RPCI-4 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP4-651E10.
FEATURES
Source
1. 113764
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="1"
   /map="p22.3-31.1"
   /clone="RP4-651E10"
   /clone_id="RPCI-4"
BASE COUNT 34890 a 21989 c 22351 g 34534 t
ORIGIN
Query Match      100.0%; Score 106; DB 9; Length 113764;
Best local Similarity 100.0%; Pred. No. 1.7e-20;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GGCATTACATTAAATAATATGATGAGAGTGGAGAACTGACGCTCAATAGCC 60
        |||||||
DB      32933 GGCATTACATTAAATAATATGATGAGAGTGGAGAACTGACGCTCAATAGCC 33012
        |||||||
QY      51 TAGGGCTGAATTTTGTGAGATAAATAATATCAATCATTCCTT 106
        |||||||
DB      33013 TAGGGCTGAATTTTGTGAGATAAATAATATCAATCATTCCTT 33058
        |||||||
RESULT 7
AL358950/c

```


LOCUS AL358950 164891 bp DNA HTG 23-JAN-2001
DEFINITION Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN
PROGRESS *** 10 unordered pieces.
ACCESSION AL358950
VERSION AL358950.4 GI:12539689
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 164891)
Plumb,B.
Direct Submission
Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
requests: clonequest@sanger.ac.uk
On Jan 26, 2001 this sequence version replaced gi:9988471.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: BA444C12
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least Q40
Consensus quality: 163328 bases at least Q30
Insert size: 163991; sum-of-ctigs
Insert size: 163373; 6.6% error; agarose-fp
Quality coverage: 5.19x in Q20 bases; sum-of-ctigs Quality
coverage: 5.21x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 34236: contig of 34236 bp in length
* 34237 34336: gap of 100 bp
* 34337 47372: contig of 13036 bp in length
* 47373 47472: gap of 100 bp
* 47473 67025: contig of 19553 bp in length
* 67026 71239: gap of 100 bp
* 71240 71339: contig of 4114 bp in length
* 71340 83903: gap of 100 bp
* 83904 84003: contig of 12564 bp in length
* 84004 91357: gap of 100 bp
* 91358 91457: contig of 7354 bp in length
* 91458 96452: gap of 100 bp
* 96453 96552: contig of 4995 bp in length
* 96553 148189: gap of 100 bp
* 148190 148289: contig of 51637 bp in length
* 148290 159778: gap of 100 bp
* 159779 159878: contig of 11489 bp in length
* 159879 164891: gap of 5013 bp in length.
FEATURES
Source
1. 164891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-444C12"
/clone_1b="RPC1-11.2"
1. 34236
/note="assembly_fragment:00027
fragment_chain:1
misc_feature
1. 34236
/note="assembly_fragment:00027
fragment_chain:1

clone_end:"7
vector_side:"left"
34337..47372
/note="assembly_fragment:00312
fragment_chain:1"
47473..67025
/note="assembly_fragment:01923
fragment_chain:1"
67126..71239
/note="assembly_fragment:02094
fragment_chain:1"
71340..83903
/note="assembly_fragment:02213
fragment_chain:1"
84004..91357
/note="assembly_fragment:01292
fragment_chain:1"
91458..96452
/note="assembly_fragment:01391
fragment_chain:1"
96553..148189
/note="assembly_fragment:02045
fragment_chain:1"
148290..159778
/note="assembly_fragment:00427
fragment_chain:1"
159879..164891
/note="assembly_fragment:00773
fragment_chain:1
clone_end:"SP6
vector_side:"right"
FASTA COUNT 48381 a 32179 c 32357 g 51067 t 907 others
ORIGIN
Query Match 100.0%; Score 106; DB 2; Length 164891;
Best Local Similarity 100.0%; Pred. No. 1.7e-20;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCATTGACATTTTAAATATGTCGATGATAGACAGACGCTCATATGCC 60
DB 49748 GGCATTGACATTTTAAATATGTCGATGATAGACAGACGCTCATATGCC 48719
QY 61 TAGGCTGATTTTGTGCAGATAAATAAATCAATCATCTT 106
DB 48718 TAGGCTGATTTTGTGCAGATAAATAAATCAATCATCTT 48673
RESULT 8
LOCUS 195746 878 bp DNA PAT 01-DEC-1998
195746
DEFINITION Sequence 8 from patent US 5733748.
ACCESSION 195746
VERSION 195746.1 GI:3940216
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 878)
AUTHORS Yu,G. and Rosen,C.
TITLES Colon specific genes and proteins
JOURNAL Patent: US 5733748-A 8 31-MAR-1998;
FEATURES
Source
1. 878
/organism="unknown"
FASTA COUNT 257 a 179 c 188 g 241 t 13 others
ORIGIN
Query Match 72.5%; Score 76.8; DB 6; Length 878;
Best Local Similarity 91.0%; Pred. No. 2.9e-12;
Matches 81; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 GGCATTGCATTTTAAATATATGAGAGTGGATGAGAGACTGCAGTGCATATGCC 60
 |||||
 Db 717 GGCATTGCATTTTAAATATATGAGAGTGGATGAGAGACTGCAGTGCATATGCC 776
 |||||
 OY 61 TAGGCTGATTTTGTGCAGATAATAA 89
 |||||
 Db 777 TAGGCTGATTTTGTGCAGATAATAA 805

RESULT 9

AF095584 3079 bp mRNA MAM 13-OCT-2000
 LOCUS Sus scrofa epithelial chloride channel protein (AECCL) mRNA,
 DEFINITION complete cds.

AF095584
 VERSION AF095584.1 GI:6002645

KEYWORDS

SOURCE

plg.
 ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

20473747

11015605

2 (bases 1 to 3079)

Gaspar, K.J., Racette, K.J., Gordon, J.R., Loewen, M.E. and Forsyth, G.W.

Cloning a chloride conductance mediator from the apical membrane of porcine ileal enterocytes

Physiol Genomics (online) 3 (2), 101-111 (2000)

Submitted (28-SEP-1998) Veterinary Physiological Sciences, University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4, Canada

FEATURES

source

Location/Qualifiers

1..3079

/organism="Sus scrofa"

/db_xref="taxon:9823"

/tissue_type="ileal mucosa"

1..3079

/gene="AECCL"

133..2886

/gene="AECCL"

/function="involved in exocrine chloride conductance"

/codon_start=1

/product="epithelial chloride channel protein"

/protein_id="AA00077.1"

/db_xref="GI:6002645"

/translation="MGSFSSFLFLLVLLHLEGAQNSLIQLNGEGYGVIAIDNPV
 EDERLIONIKDWTASPYLFEATEKREFKFNVALILIPASMAKPEYVPLKLETKNA
 DVAVTEPNPNDGPTTEBONGGEGEYVTPDFVAGKVLQYGPQGRVHVMAM
 LRMGVFENYNNOKFYLSNKKRQVYCSAIRTNLVPCOGGSCVTKRCADRTGL
 FQKCEFTIDPOOSKASTMPSIDTYVEPKENHNKAEADNCONKCNSTMEVI
 DQSEDEKTTPTPTDPPATFTSLDQIGATVLYVDKSSMTVSGRKLRLNAGTLFL
 LQVEGAGWAVAFDSDAAYVSELYQVISAERDLARSLTPASGGTISCSGLRSA
 FVIRKKYPTDSEIVLLTDGDNLTISAFPEVKONGALITHTVALGSAKLLELSQ
 MTGGLOTYASDQENNGLIDAFGALSNGRAASQSIQLESGLTQNNENMGTVV
 DSTVGDFTLTLLEKFLSPIPEFGVGRSDSFLVGHKMAFYQVGRATKQVM
 KYSAQSOTLTLTVSSRSASATLPVTVTSKMNKDTGKSPVAVYTKIHGTLPIIL
 RAKVTALISENGKTYTLLDNGAGADATKNDGYSRFTAYDANGRSVAVMLGG
 VMTPRRAPPVLSGMYTIRGWIENGELIKNNPRPDINKDDLGKQVCSGGSFV
 ASDVPSKPIPLDFPPCKITDLKAGIAGDMILNLTWAPNGDDHGRADRYLIRISNI
 LDIKDFNDVQVNTDILPKKANSEEVVFPDEGIPFTNGTDLFLAVQAVKTNLKS
 EISNIAQVSLPLPEAPPEPTPTAPSLPCEIQLVNSTIPGIIHILIMKMLGELQL
 STA"

BASE COUNT 892 a 758 c 755 g 674 t

ORIGIN

Query Match 67.4%; Score 71.4; DB 4; Length 3079;
 Best Local Similarity 87.6%; Pred. No. 1.2e-10;
 Matches 78; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 GGCATTGCATTTTAAATATATGAGAGTGGATGAGAGACTGCAGTGCATATGCC 60
 |||||
 Db 2824 GGCATTGCATTTTAAATATATGAGAGTGGATGAGAGACTGCAGTGCATATGCC 2883
 |||||
 OY 61 TAGGCTGATTTTGTGCAGATAATAA 89
 |||||
 Db 2884 TAGGCTGATTTTGTGCAGATAATAA 2912

RESULT 10

AB017156 2937 bp mRNA ROD 10-NOV-1999
 LOCUS Mus musculus gob-5 mRNA, complete cds.

AB017156
 VERSION AB017156.1 GI:3721911

KEYWORDS

SOURCE

Mus musculus adult intestine goblet cell cDNA to mRNA.
 ORGANISM

AUTHORS

TITLE

JOURNAL

99160866

2 (bases 1 to 2937)

Komiyama, T., Tanigawa, Y. and Hirohashi, S.

Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice

Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)

Submitted (26-AUG-1998) to the DDBJ/EMBL/GenBank databases, Tohru Komiyama, ERATO, JST, Genome Asymmetry Group, DAI Biosymmetry Project, Tokodai 5-9-9, Tsukuba, Ibaragi 300-2635, Japan (E-mail:tkom@bioa.eriato.ttc-net.co.jp, Tel:81-298-48-1515, Fax:81-298-47-8901)

FEATURES

source

Location/Qualifiers

1..2937

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_type="goblet cell"

/dev_stage="adult"

15..2756

/tissue_type="intestine"

15..2756

/gene="gob-5"

15..2756

/gene="gob-5"

/function="gob-5"

/codon_start=1

/protein_id="BAA3743.1"

/db_xref="GI:3721912"

BASE COUNT 860 a 718 c 693 g 666 t

ORIGIN

Query Match 44.7%; Score 47.4; DB 10; Length 2937;
 Best Local Similarity 70.8%; Pred. No. 0.00094;
 Matches 63; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

SOURCE
ORGANISM
Bacteria: Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
1 (bases 1 to 29324)
Seanan, E. and Devine, K.M.
Sequence of the Bacillus subtilis chromosome from ykua to cse-15
Unpublished
2 (bases 1 to 29324)
Devine, K.M.
Direct Submission
Submitted (06-NOV-1997) K.M. Devine, Genetics, University of
Dublin, Trinity College, Lincoln Place Gate, Dublin 2, Ireland
Location/Qualifiers
1. 29324
/organism="Bacillus subtilis"
/strain="168"
/db_xref="taxon:1423"
89. .955
/gene="ykc"
89. .955
/gene="ykc"
/note="homologous to 3-hydroxyisobutyrate dehydrogenase"
/codon_start=1
/transl_table=11
/product="ykc protein"
/protein_id="CAI0859.1"
/db_xref="GI:2632217"
/db_xref="SWISS-PROT:O34948"
/translation="MKTTIGFISGVGKSMASHILNDHPVLVYRTKREKASILQK
GAIMKDVXDLSEADVITIMVGPSPVEEYVSGNIIENAKGAILDITMSKPEL
AKTIAEKAKALFALDAPYGGIDNGSLAIMWGEEKAFECACWILITMSKPEL
YOGPAGSGOHTKMCNOAIIAGMIGVGAEMAAVYAKSGLEPENVLKSTTGAAGSMIS
NLAPRMQGNFEPGFYKHFIKMGIALFELALMGEMPCISAKSLYDKLAAOGSEN
SGOSITKLVNK"
complement(994. .1767)
/gene="ykw"
/complement(994. .1767)
/gene="ykw"
/codon_start=1
/transl_table=11
/product="ykw protein"
/protein_id="CAI0860.1"
/db_xref="GI:2632218"
/db_xref="SPRMBL:O31398"
/translation="MKKAPILSAAAVGLFTFGVQOASAKELSCOPVTVTKGNTVO
NMSLNDVAKLHNTNITLNAANEKEMKQLQKHAQSNVQVOVKETAPPAKT
TERAADQONTASAPATAEKNTTTSAPSSVSAVERKVVETLNAEROKLQLOIDE
TLSKSAKASQDMKKNFDHOSPTYSPPDMAKSGISYKTAGEINIAKQKPEEV
KAMNSEGHKRLNPNFTIIGVYESSGIWQOFIGK"
2104. .4218
/gene="yku"
2104. .4218
/gene="yku"
/note="homologous to penicillin binding protein 2b"
/codon_start=1
/transl_table=11
/product="yku protein"
/protein_id="CAI0861.1"
/db_xref="GI:2632219"
/db_xref="SPRMBL:O31399"
/translation="MTEIGREPKSKSGNRAIRBNLFPLAVFVLTALIFKLGVQIV
EGROHEDAEKSNATAYYPAPRGKMDROKAYVNOASPEIYVSTSTKREDKIK
TAKRLASFTHIDTEFLAKERLDPIYTAHKAALIKSESNLKGQAKLIDERNP
DOFLKAIQDDDEEMETAIYTRFSSGNAEPQVAKANPKSNSNGKNGALDEKNS
SOPKNDLTYDELSIVSEHLELPGIDIVDMTRKPYDKTILSVFGVTPPOGLIS
DRKDFLTLYGANNDRVKSLEYOEYEEYINSHKEVEYEDKNGNVSOITDKGSR
GYLOLSPMELOAKVEKIIIEEVNSRANGVMDRAFAVMMDPNNNGDLSMAKKI
DKTNKIEDYATGALTTOYEMGSAVGATVLAGYODGIPHYKYITADAPMLGNNLIK
STNMGTINELTALOKSSNVYMNVMHAGVTYKPHGSLPADQNDLAKRNTYSOPG

LGKVTGIDLPQESAGMOTETKTVGGLILDLAIGQYDYTYELQANAQYISVANGYRVQ
PRIVTSIHKPKDKOLGKALEHRKPKVLNKNINSESDLOQVOTGMKLVTSSTAKNF
TEDVSGTGTAEFFYYGTNNRMNGKRTYNLFVGYYSKKPKVAFSVVPVDDDKI
NKTIARAIHAYAELEKKHSK"
4382. .6202
/gene="kin"
4382. .6202
/gene="kin"
/function="sporulation kinase"
/codon_start=1
/transl_table=11
/product="kin protein"
/protein_id="CAI0862.1"
/db_xref="GI:2632220"
/db_xref="SWISS-PROT:P16497"
/translation="MEQDQHVHPLQKTHQHVIVASNGRIIYASNRKHLGYGE
MISFLKTFHEEDQPLYESYFNEHNMCTREPIKDDTIYWEANAENVITTRAR
TERITILKMYLEBETHOSLCKEHEITEPASPSYITIDYERLENPSPICISV
KGIIVYNSAMLSMLAKSKDAIIGKSSYFIEEYHDIYKNIIRKQKMGEMIEQ
TKRLDQTPVHLEKASPIYKNOQALDILIDISSKKQTLQKLSREYOLLIONS
IDTIAVHNGKWFVNMESGISLEAATYEDLIGNIYDOLHPCDHEDVKERIONIAQ
KTESEIYKOSMFTFONRVITYTEWVCIPTEFGEAAVYILDIISERKOTELKISK
LSIAGQAAAGIAHRIIRPLAIGFIQIMKPYMEGNEHYDIYFSELSREILSEL
MLAKPOONAVEKYLNLAKLIGVYSALETOANGLIFRTYERKSDTYINGDONQKQ
VFNLINKNAVESHPDGGTDVLIITEHSHVAVYKDEGCIPEKRYLNRIGEPLITKE
KGTGLMVFNTIENHOGVYIHWDSHEKTAFKISPPK"
complement(6213. .7391)
/gene="pata"
complement(6213. .7391)
/gene="pata"
/note="hisC homologue"
/codon_start=1
/transl_table=11
/product="pata protein"
/protein_id="CAI0863.1"
/db_xref="GI:2632221"
/db_xref="SWISS-PROT:P16524"
/translation="MEHLNPKAREIEISGIRKFSNLVQHEDEVISLTIGQDPFPH
HYKAAKKAIDENVTSTPNAGYLELROAVOLMKKADENYAESEIIITTCOAID
AARFLLSPGEYIMGPPIYGVGEPIILNCGARVYDPTSHGFKTLARLIEALPN
TKCVLPYPSPNGCVTLSEELKSTALILKGRVYPLSDIYSLEYTRHYSIATYL
RDQTIYVNGISKSHSMTGRIGFLFAKADIAKHLVYHOVNCASSISOKALEAVT
NGLDALIMREQKRLDYDRLVSMGLDVPVPSAFYIFPSIKSGMISPDFSMAL
LEDQVALVPQSSFSYEGGYVRLSPFASMDTLREGLDLRELVKREMQVTINNGV"
7958. .8869
/gene="chev"
7958. .8869
/gene="chev"
/function="Involved in chemotaxis"
/codon_start=1
/transl_table=11
/product="Chev protein"
/protein_id="CAI0864.1"
/db_xref="GI:2632222"
/db_xref="SWISS-PROT:P37599"
/translation="MSLQYEILDSGTNELIVKFGVGNARGINVMKVRILIOPE
VTSVPSHSHOVECMIKRLGELLPLVISLFSFGVEPSSDDEKTYIYEPKRRKYFVQ
SVQIHRYSNEALEKPTSLNOGEMERHITGIKLEDMITFLPYEKIITYIDESGDT
YNNHTGFDERKDKLITVEDSPLMLRLQDELKAGYNNINLAFENGAEAYINML
AENETDLSKQIDMIYDIDEMPKMDGRHLRLKLKENDKSDVPVMPFSSLTIDDLRRHG
EVVGADEQISKEPISDLIKVDYVYE"
complement(8913. .9377)
/gene="ykuB"
complement(8913. .9377)
/gene="ykuB"
/codon_start=1
/transl_table=11
/product="ykuB protein"
/protein_id="CAI0865.1"
/db_xref="GI:2632223"
/db_xref="SPRMBL:O34548"
/translation="MDHAYTRKDLQPTVENLSKAVYTVNRHAKTAPNPKYLYLLKKRA


```

LOCIVEGKGGKKGICGHEKSNPREFSOQSDVLISIDYVFNHPEPTDEDEHLEHFLCTLN
QSYRNPKAQMSLTKAKHLLOEYVGRKREPLVNPQAPYHKKFVKLESTYF"
gene
  complement(9503. .10795)
  /gene="yKuc"
  complement(9503. .10795)
  /note="homologous to macrofilide efflux determinants"
  /codon_start=1
  /transl_table=11
  /product="YKuc protein"
  /protein_id="CA10866.1"
  /db_xref="GI:2632224"
  /db_xref="SPTREMBL:O31400"
  /translation="MDIFKNRNFVRLFFAALASQMGTTVGNNAFAPFLIDRFSQPSY
  TLAELMSLPTEFVFLVGVADRFDRKVAENCDMIRAGLTVLVLPFLPGNIPLV
  FCLIEBSAVTKFFEPANSLVOALTPKHEAKAGLNMPLSFVGVGIGTGAEMV
  TIGIEGATILDPVPSFTISGLIRSCRIPKEAKOPNGARSMKRTSKDSNDREGLY
  ILKKNLKLSTLFFFTFIFGVNGFVAVLPMTKTKGLADPRTKEMTSVFALGFGELA
  GSVIGTILSKRVKPKPELMSIPFVAGLLIFVLYGYNILMYYVAAAFALGMCIGPINA
  IGWMPKIVHPRKMGVSGMODPEMFAOSLIGLVALFPFVSNIDYLYXGKGVII
  LLVFFNFIALPKYSQAQAEVAVCEAFQEPQPKKRSY"
gene
  complement(10871. .11365)
  /gene="yKud"
  complement(10871. .11365)
  /gene="yKud"
  /codon_start=1
  /transl_table=11
  /product="YKud protein"
  /protein_id="CA10867.1"
  /db_xref="GI:2632225"
  /db_xref="SPTREMBL:O34816"
  /translation="MLTYOVKOGDTLNSIADFRISTALLQANPSLQAGTLAGOSIV
  IPGADPYTIPVHIVSIGAKTLTSLNNRVAKGPIAVGKILVOTPTGCEFTINROR
  NPGCGPAGVWVLSLQKHGIGHTNPNPASIGKAVSGCIRMHKKDVIELASIVNCRIV
  TINR"
gene
  complement(11422. .12285)
  /gene="yKue"
  complement(11422. .12285)
  /gene="yKue"
  /codon_start=1
  /transl_table=11
  /product="YKue protein"
  /protein_id="CA10868.1"
  /db_xref="GI:2632226"
  /db_xref="SPTREMBL:O34870"
  /translation="MKKMSRQFLKMGALAGALTAAGCGYARLEPHMTEETH
  TIKSLIPHGDFGKIVQPSDAHSDYFTLELKVYIILTINSKEDLVYFTGDIIDNP
  DYOHHQAVIPLLRLNAPFGKLCITYGNHHDGCGTAVYKSLMTAGGFTVYRNGYQTL

```

Query Match 34.5%: Score 36.6; DB 1; Length 29324;
 Best Local Similarity 60.6%: Pcore No. 1.5;
 Matches 60; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```

QY 1 GGCATTGACATTTTAAATATATGGAAGTGAAGAGAACTGACGTGCAATAGCC 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7888 GCCATTCACCTTTTACAAATGATATACGCGATATACGATGCAATGATGCG 7947
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 TAGGCGCTGATTTTGTGCAATTAATAAATAATCAAT 99
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7948 AGGAGTGAATTGCTGTACAAACATACGAAATTTATT 7986

```

RESULT 14
 LOCUS BSUB0008 208230 bp DNA BCT 26-NOV-1997
 DEFINITION *Bacillus subtilis* complete genome (section 8 of 21): from 1394791
 to 1603020.
 ACCESSION Z99111 AL009126
 VERSION Z99111.1 GI:2633699
 KEYWORDS
 SOURCE *Bacillus subtilis*.
 ORGANISM *Bacillus subtilis*
 Bacteria; Firmicutes; *Bacillus*/Clostridium group;

```

REFERENCE
  1 (bases 1 to 208230)
  Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,
  Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S.,
  Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C.,
  Bron, S., Brouillet, S., Brusch, C.V., Caldwell, B., Capiano, V.,
  Carter, N.M., Choi, S.K., Codani, J.J., Connerion, I.F., Cummings, N.J.,
  Daniel, R.A., Denicola, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,
  Emerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E.,
  Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Funa, S., Gallizzi, A.,
  Galleron, N., Ghim, S.T., Glaser, P., Gottreau, A., Gollightly, E.J.,
  Grandi, G., Guisepi, G., Guy, B.J., Haga, K., Halech, J., Harwood, C.R.,
  Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hulio, M.F.,
  Itaya, M., Jones, L., Joris, B., Karmata, D., Kasahara, Y.,
  Klaer, Blanchard, M., Klein, C., Kobayashi, Y., Koeltter, P.,
  Koningsstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,
  Lathinols, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H.,
  Masuda, S., Maue, C., Medigue, C., Medina, N., Mellado, R.P.,
  Mizuno, M., Mostl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,
  Ogawa, K., Ogilwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M.,
  Portetle, D., Portwoll, S., Prescott, A.M., Prescan, E., Pujic, P.,
  Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M.,
  Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadale, Y., Sato, T.,
  Scallan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,
  Sekowska, A., Seror, S.J., Serro, P., Shin, B.S., Soldo, B.,
  Sorokin, A., Taccioni, E., Takagi, T., Takahashi, H., Takemaru, K.,
  Takeuchi, M., Tamakoshi, A., Tanaka, T., Tempstra, P., Togonoi, A.,
  Tostato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A.,
  Viari, A., Wambolt, R., Wedler, E., Wedler, H., Weitzel, G., Yata, K.,
  Winters, P., Wipal, A., Yamamoto, H., Yamane, K., Yasunoto, K., Yata, K.,
  Yoshida, K., Yoshikawa, H.F., Zumbstein, E., Yoshikawa, H. and
  Danchin, A.
  The complete genome sequence of the gram-positive bacterium
  Bacillus subtilis
  Nature 390 (6657), 249-256 (1997)
  98044033
  2 (bases 1 to 208230)
  Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.
  Direct Submission
  Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur,
  Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724
  Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
  adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
  68 89 48
  FEATURE
  source
  1.208230 Location/Qualifiers
  /organism="Bacillus subtilis"
  /strain="168"
  /db_xref="taxon:1423"
  46.183
  /gene="yKZD"
  46.183
  /gene="yKZD"
  /function="unknown"
  /note="similar to hypothetical proteins from B. subtilis"
  /codon_start=1
  /transl_table=11
  /protein_id="CAB13202.1"
  /db_xref="GI:2633700"
  /db_xref="SPTREMBL:O34405"
  /translation="MEKKEQYINQAEVYVPHPTKRGVAFVLFHETVHLSEDDTQTT
  E"
  689.2044
  /gene="yKOK"
  689.2044
  /gene="yKOK"
  /function="unknown"
  /note="similar to Mg2+ transporter"
  /codon_start=1
  /transl_table=11
  /protein_id="CAB13203.1"
  /db_xref="GI:2633701"
  /db_xref="SPTREMBL:O34442"

```


[illegible]

gene
CDS
/gene="BB0092".
complement(3430. .4044)
/note="similar to GB:U77117 SP:Q58032 PID:1592298 percent
identity: 26.24; identified by sequence similarity:
putative"
/codon_start=1
/transl_table=11
/product="V-type ATPase, subunit D (atpd)"
/protein_id="AAC66485.1"
/db_xref="gi:2687981"
/translation="MIMSKIKILKTNLKKDELMKFRRLPTQLKKQOLYMEIVAKK
ENSVYKMLDEQOKLENISSMLSEKFEPSMIOYTVKSSINAGVAPIEDFSI
EVEDRHLDELPTPWMDKGIILNVIOIDVELIKKIDOLLREPRITSQRANLEF
KVMIPAAINAKIKINIYIGDOOTAAVVRGKIAKSLIKKNNSL"
complement(4041. .5345)
/gene="BB0093"
complement(4041. .5345)
/gene="BB0093"
/note="similar to GB:J04836 SP:P22663 PID:149821 percent
identity: 38.53; identified by sequence similarity:
putative"
/codon_start=1
/transl_table=11
/product="V-type ATPase, subunit B (atpb)"
/protein_id="AAC66484.1"
/db_xref="gi:2687980"
/translation="MKRYSKIEISAGVITYTNOGIRYGLAIVAKKADTSSLAETIKR
LDREAVSIAQVGGIRGVSTDEIKIFLGHSQVVSFSDNLGIFPGSGNPRDGGSLDD
NLIEIGSGASNPQKIVPRNMIKRGLEPIADVENLVESQKLPITSVSGEPNELLIR
ALQAEVILITGKMKLKHDDYLTFKDSLEKGLASRAIFVFANDSVESLVLTIR
LSVAEKFALCKKIVYLTDITNFDADAKKEISITMEQVPSNRGPGDLYSGLAVRE
AIDFEGASITILITVTMPGDVTHPYPDNNTGYITEGOYLYKGRIEPGLSLKKOM
VNSRRDHRITMDSMTIKLVASKESEYKRAMGFMTRKMDKILKISNMFESKMDLS
VNIPLEELDLGKSLASCFSPEKTEGITDLIEKRWKKET"
complement(5367. .7022)
/gene="BB0094"
complement(5367. .7022)
/gene="BB0094"
/note="similar to GB:J04836 SP:P22662 PID:149820 percent
identity: 46.96; identified by sequence similarity:
putative"
/codon_start=1
/transl_table=11
/product="V-type ATPase, subunit A (atpa)"
/protein_id="AAC66483.1"
/db_xref="gi:2687979"
/translation="MNEVLPVKTAGRNLKAVIRIRGNENAOVPELTKGISVDIYR
PTDRLITVELDGLITVOYDGLQNPRLPELAIQCGFELERGVLRPLNKDKWNRKRTS
KVGDIIVAGDELGVETEGTVHQIMLIPYKRDSTYVEIVSDGYSTDEQIAVLIEDSS
GMRNITISFMIPVQVPIPTNKERLISEPEMLTQTRITIDFFPAKGFCTCIPGFGA
GKTIVAOQTSNADVDVILAACGERAAVEVETLEKEPDLADPTGKSIMDRCTIEN
TSSMVAAREASVYTAIRIGEYROMGDDILLADSTSRMAOARMKSGRALEIDGEE
APALTEVASISFEKRAIGIVLWNGDIGSVVGGSPAGCNPREPYQALTLKYVGA
HGURPERDARKPAISPLESMKRYGTYTDKRTKYRASFVLYKNEINOMAKVYGBEG
ISNDFELIYLSLSDSCYLQDNSSTIDAAVSSRONVMTDVIINILTKTFESDKIK
QARPEILNLRNLLDMNLSPKDNHFKNLEHALGELINFKKVI"
complement(7108. .7653)
/gene="BB0095"
complement(7108. .7653)
/gene="BB0095"
/note="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB0095"
/gene="BB0095"


```
/protein_id="AAC66489.1"
/db_xref="GI:2687985"
/translation="MLDSYYVLSLPYIDLKLKNSVSDFLNWEISLTKDENFL
KELLEFKIDRGSRVIDLFLNFEDTIRVTLAALRAKLGSKDYLESTYSSYLSI
LKNICLKENPEFEIISDLKQFLDLEVGHEFDEKLIYFLKRLFLRNLPEKEK
LGIONFENICKNLIDKTNEKV"
complement(7663..8265)
/gene="BB0096"
/complement(7663..8265)
/gene="BB0096"
/note="similar to GB:I77117 SP:Q57673 PID:1590957 percent
identity: 27.69; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=1
/product="V-type ATPase, subunit E, putative"
/protein_id="AAC66482.1"
/db_xref="GI:2687978"
/translation="MMQFEVKDINKIKKDGLEAEERASNDIILKAKREAEIVARAE
EAARVLAKSEKENDYKCHALEASROATLITGVEKNLSEFNALKDNVEVLSL
NNELAEILIKITDSMAKGEKIVIOLESDSLSLQIIRLKGKRLKGMELRPPRCIS
KGFRIQKNIGLHYDESSETIADLFYDLNPRREVITKV"
complement(8449..8835)
/gene="BB0097"
/complement(8449..8835)
/gene="BB0097"
/complement(8449..8835)
/gene="BB0097"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BB0097"
/protein_id="AAC66488.1"
/db_xref="GI:2687984"
/translation="MNTSKFNEFEENVLFHICLDVDFVLENEFGNSYDIHPNRPGRKA
ANGLLDGLFSVTTFTSGSGSRFGRGVIILITLNLFPVDFEDKINKKQIELFRSG
LKNKFPKNDIVLDGNVYKIIGPF"
complement(8841..11183)
/gene="BB0098"
/complement(8841..11183)
/gene="BB0098"
/complement(8841..11183)
/note="similar to PID:1652751 percent identity: 28.43;
identified by sequence similarity; putative"

gene
CDS

Query Match
Best Local Similarity 59.2%; Pred. No. 7.7;
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
Score 34; DB 1; Length 12984;

OY 5 TTCACATTTTAAATATATGGAAGTGAAGAGAGACTGCATGACCTAGG 64
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2229 TACAAAATATATCTTATCTGTTTCAATAGGGGTTTGCAAAATTCATTGCGCATG 2170
| | | | | | | | | | | | | | | | | | | | | | | | | |
OY 65 GCTGAATTTTGCAGATAAATAAATAATCATTCAT 102
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2169 TTTGGAAATTTTCAGACAGTAAGAAAGACCCCAT 2132
| | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: April 4, 2002, 03:08:35
Job time: 60797 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:31 ; Search time 16681 Seconds
(Without alignments)
147.520 Million cell updates/sec

Title: US-09-049-696-10
Perfect score: 229
Sequence: 1 GGCACTGATGCTGGACAG.....AACCTGACCTGACTGTCA 229

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 segs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estto:*
8: em_estov:*
9: em_hic:*
10: gp_est1:*
11: gp_est2:*
12: gp_hic:*
13: gp_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	229	100.0	625	10	AW361532 QV2-CT026
2	227.4	99.3	653	10	AW361534 QV2-CT026
3	225.8	98.6	561	10	AW361523 QV2-CT026
4	224.2	97.9	621	10	AW361520 QV2-CT026
5	224.2	97.9	633	10	AW753451 QV2-CT026
6	224.2	97.9	652	10	AW753449 QV2-CT026
7	222.6	97.2	657	10	AW361521 QV2-CT026
8	200.8	87.7	356	10	AW361528 QV2-CT026
9	170.6	74.5	654	10	AW361522 QV2-CT026
10	161.8	70.7	455	10	AA871197 VQ32C11.1
11	161.8	70.7	2915	12	AA8007466 Mus muscu
12	161.8	70.7	2933	12	AK008659 Mus muscu

13	151.4	66.1	906	11	BF581342
14	151.4	66.1	935	11	BF578919
15	139	60.7	496	10	AW753448 QV2-CT026
16	138.2	60.3	308	10	AV077718 AV077718
17	137.2	59.9	469	10	BE136935
18	107.4	46.9	550	11	BE355533
19	105.8	46.2	534	11	BE355513
20	102.2	44.6	974	11	BE355513
21	101.6	44.4	661	11	BE355513
22	83.8	36.6	1107	11	BF581041
23	76.2	33.3	911	11	BF581120
24	74	32.3	747	11	BF580957
25	73	31.9	742	13	BH030431
26	63	27.5	479	10	AA541829
27	60	26.2	549	10	BE281945
28	59.4	25.9	235	10	BB299286
29	59	25.8	510	10	AW762590
30	59	25.8	866	11	BI155803
31	59	25.8	914	11	BG172513
32	58.2	25.4	584	11	BF116637
33	56.4	24.6	796	11	BI157217
34	54	23.6	583	10	AW230825
35	54	23.6	625	10	AW230824
36	51.2	22.4	142	11	BE333499
37	50.4	22.0	301	13	AZ779200
38	50	21.8	1033	11	BG243305
39	45.6	19.9	462	10	A1851265
40	44.6	19.5	279	10	BB475477
41	43.6	19.0	299	10	BB124752
42	43.2	18.9	266	10	BB072609
43	42.6	18.6	539	13	A2464851
44	42.6	18.6	703	13	A2022011
45	42.4	18.5	300	10	BB560274

ALIGNMENTS

RESULT 1
AW361532 625 bp mRNA EST 04-FEB-2000
DEFINITION QV2-CT0261-261099-011-e03 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW361532
VERSION AW361532.1 GI:6866286
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 625)
AUTHORS HCCP <http://www.ludwig.org.br/ORESTES>.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil.
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL.
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&ct=QV2-CT0261-261099-011-e03&tl=1999-10-26&tl=1>)
Seq primer: puc 18 forward
High quality sequence start: 93
High quality sequence stop: 608.
Location/Qualifiers
1..625
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"

FEATURES

source

/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 172 a 154 c 150 g 149 t
ORIGIN

Query Match 100.0%; Score 229; DB 10; Length 625;
Best Local Similarity 100.0%; Pred. No. 4.2e-60;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACAGTATCTGTGACAGCACCCTGGGAAAGACACTTTGTTTATCAGCTGGACA 60
DB 360 GGCACAGTATCTGTGACAGCACCCTGGGAAAGACACTTTGTTTATCAGCTGGACA 419
QY 61 ACGCAGCTCTCCCAAAATCTTCTGTGGATCCAGTGGACAGAGAGAGTGGCTTGT 120
DB 420 ACGCAGCTCTCCCAAAATCTTCTGTGGATCCAGTGGACAGAGAGAGTGGCTTGT 479
QY 121 GTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTAAGTTGGCACT 180
DB 480 GTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTAAGTTGGCACT 539
QY 181 TGGAAATACAGTCTGCAGACAGCTCACAACCTTGACCTGACTGCA 229
DB 540 TGGAAATACAGTCTGCAGACAGCTCACAACCTTGACCTGACTGCA 588

RESULT 2
AM361534 653 bp mRNA EST 04-FEB-2000
LOCUS QV2-CT0261-261099-011-f01 CT0261 Homo sapiens cDNA, mRNA sequence.
DEFINITION AM361534
ACCESSION AM361534.1 GI:6866288
VERSION
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 653)
HCCP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?fil=QV2&f2=QV2-CT0261-261099-011-f01&f3=1999-10-26&f4=1)
Seq primer: puc 18 forward
High quality sequence start: 37
High quality sequence stop: 652.
Location/Qualifiers
1. .653

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 154 a 140 c 140 g 127 t
ORIGIN

into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 140 a 156 c 166 g 191 t
ORIGIN

Query Match 99.3%; Score 227.4; DB 10; Length 653;
Best Local Similarity 99.6%; Pred. No. 1.3e-59;
Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCACAGTATCTGTGACAGCACCCTGGGAAAGACACTTTGTTTATCAGCTGGACA 60
DB 628 GGCACAGTATCTGTGACAGCACCCTGGGAAAGACACTTTGTTTATCAGCTGGACA 569
QY 61 ACGCAGCTCTCCCAAAATCTTCTGTGGATCCAGTGGACAGAGAGAGTGGCTTGT 120
DB 568 ACGCAGCTCTCCCAAAATCTTCTGTGGATCCAGTGGACAGAGAGAGTGGCTTGT 509
QY 121 GTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTAAGTTGGCACT 180
DB 508 GTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTAAGTTGGCACT 449
QY 181 TGGAAATACAGTCTGCAGACAGCTCACAACCTTGACCTGACTGCA 229
DB 448 TGGAAATACAGTCTGCAGACAGCTCACAACCTTGACCTGACTGCA 400

RESULT 3
AM361523 561 bp mRNA EST 04-FEB-2000
LOCUS QV2-CT0261-261099-011-g05 CT0261 Homo sapiens cDNA, mRNA sequence.
DEFINITION AM361523
ACCESSION AM361523.1 GI:6866277
VERSION
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 561)
HCCP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?fil=QV2&f2=QV2-CT0261-261099-011-g05&f3=1999-10-26&f4=1)
Seq primer: puc 18 forward
High quality sequence start: 43
High quality sequence stop: 560.
Location/Qualifiers
1. .561

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 154 a 140 c 140 g 127 t
ORIGIN

Query Match 98.6%; Score 225.8; DB 10; Length 561;
 Best Local Similarity 99.1%; Pred. No. 4e-59;
 Matches 227; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGCACAGTATCGTGGACAGCACCCTGGGAAAGACACTTTGTTCTTATCACCCTGGACA 60
 |||||
 Db 314 GGCACAGTATCGTGGACAGCACCCTGGGAAAGACACTTTGTTCTTATCACCCTGGACA 373
 QY 61 AGCAGGCTCCCAAAATCCTCTCTGGGATCCAGTGAGACAGCAAGGAGGCTTTGTA 120
 |||||
 Db 374 AGCAGGCTCCCAAAATCCTCTCTGGGATCCAGTGAGACAGCAAGGAGGCTTTGTA 433
 QY 121 GTGACAAAAACACCAAAATGAGCTTACCTCAATCCAGGATTTGTAAGTTGGCACT 180
 |||||
 Db 434 GTGACAAAAACACCAAAATGAGCTTACCTCAATCCAGGATTTGTAAGTTGGCACT 493
 QY 181 TGGAAATACAGTCTGCAAGCAAGCTCAAAACCTTGACCTGACTGCA 229
 |||||
 Db 494 TGGAAATACAGTCTGCAAGCAAGCTCAAAACCTTGACCTGACTGCA 542

RESULT 4
 AM361520 621 bp mRNA EST 04-FEB-2000
 LOCUS OY2-CT0261-261099-011-e07 CT0261 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM361520
 ACCESSION AM361520.1 GI:6866274
 VERSION EST.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 621)
 HCGP <http://www.ludwig.org.br/ORESTES>.
 TITLE The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV2&ct2-QV2-CT0261-261099-011-e07&ct3-1999-10-26&ct4-1>)
 Seq primer: puc 18 forward
 High quality sequence start: 88
 High quality sequence stop: 613.

FEATURES
 Location/Qualifiers

1..621
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0261"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 166 a 151 c 150 g 154 t
 ORIGIN

Query Match 97.9%; Score 224.2; DB 10; Length 621;
 Best Local Similarity 98.7%; Pred. No. 1.3e-58;
 Matches 226; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 21 GGCACAGTATCGTGGACAGCACCCTGGGAAAGACACTTTGTTCTTATCACCCTGGACA 60
 |||||
 Db 358 GGCACAGTATCGTGGACAGCACCCTGGGAAAGACACTTTGTTCTTATCACCCTGGACA 417
 QY 61 AGCAGGCTCCCAAAATCCTCTCTGGGATCCAGTGAGACAGCAAGGAGGCTTTGTA 120
 |||||
 Db 418 AGCAGGCTCCCAAAATCCTCTCTGGGATCCAGTGAGACAGCAAGGAGGCTTTGTA 477
 QY 121 GTGACAAAAACACCAAAATGAGCTTACCTCAATCCAGGATTTGTAAGTTGGCACT 180
 |||||
 Db 478 GTGACAAAAACACCAAAATGAGCTTACCTCAATCCAGGATTTGTAAGTTGGCACT 537
 QY 181 TGGAAATACAGTCTGCAAGCAAGCTCAAAACCTTGACCTGACTGCA 229
 |||||
 Db 548 TGGAAATACAGTCTGCAAGCAAGCTCAAAACCTTGACCTGACTGCA 586

RESULT 5
 AM753451 633 bp mRNA EST 28-APR-2000
 LOCUS OY2-CT0261-261099-011-d05 CT0261 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM753451
 ACCESSION AM753451
 VERSION AM753451.1 GI:7668383
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 633)
 Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.G.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL MEDLINE
 2020263
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV2&ct2-QV2-CT0261-261099-011-d05&ct3-1999-10-26&ct4-1>)
 Seq primer: puc 18 forward
 High quality sequence start: 66
 High quality sequence stop: 90.

FEATURES
 Location/Qualifiers

1..633
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0261"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 136 a 151 c 161 g 185 t
 ORIGIN

Query Match 97.9%; Score 224.2; DB 10; Length 633;

Best Local Similarity 98.7%; Pred. No. 1.3e-58;
Matches 226; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGCACAGTATGCTGGACAGACCGTGGGAAGACACTTTGTTTATCACCCTGGACA 60
DB 573 GGCACAGTATGCTGGACAGACCGTGGGAAGACACTTTGTTTATCACCCTGGACA 514
OY 61 AGGACAGCTCCCAATCTCTCTGGGATCCAGTGGACAGAGAGTGGCTTTGTA 120
DB 513 ATGACAGCTCCCAATCTCTCTGGGATCCAGTGGACAGAGAGTGGCTTTGTA 454
OY 121 GTGGACAAACACCAAAATGGCTTACCTCAATCCAGGCAATGCTTAAGTTGGCACT 180
DB 453 GTGGACAAACACCAAAATGGCTTACCTCAATCCAGGCAATGCTTAAGTTGGCACT 394
OY 181 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGACCTGACTGCA 229
DB 393 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGACCTGACTGCA 345

RESULT 6

LOCUS AW753449 652 bp mRNA EST 28-APR-2000
DEFINITION QV2-CT0261-261099-011-e08 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW753449
VERSION AW753449.1 GI:7668381
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 652)
Dias Neto, E., Garcia Correia, R., Verjovski-Almeida, S., Britones, M. R.,
Nagai, M. A., da Silva, W. J., Zago, M. A., Bordin, S., Costa, F. F.,
Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,
Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare,
M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and
Simpson, A. J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
COMMENT 20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1-QV2&l2-QV2-CT0261-
261099-011-f03&l3-1999-10-26&l4-1)
Seq primer: puc 18 forward
High quality sequence stop: 14.

FEATURES

source 1. 652
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site.1: Sma1; Site.2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 179 a 153 c 161 g 159 t
ORIGIN

Query Match 97.9%; Score 224.2; DB 10; Length 652;
Best Local Similarity 98.7%; Pred. No. 1.3e-58;
Matches 226; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGCACAGTATGCTGGACAGACCGTGGGAAGACACTTTGTTTATCACCCTGGACA 60
DB 478 GGCACAGTATGCTGGACAGACCGTGGGAAGACACTTTGTTTATCACCCTGGACA 477
OY 61 AGGACAGCTCCCAATCTCTCTGGGATCCAGTGGACAGAGAGTGGCTTTGTA 120
DB 478 ATGACAGCTCCCAATCTCTCTGGGATCCAGTGGACAGAGAGTGGCTTTGTA 537
OY 121 GTGGACAAACACCAAAATGGCTTACCTCAATCCAGGCAATGCTTAAGTTGGCACT 180
DB 538 GTGGACAAACACCAAAATGGCTTACCTCAATCCAGGCAATGCTTAAGTTGGCACT 597
OY 181 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGACCTGACTGCA 229
DB 598 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGACCTGACTGCA 646

RESULT 7

LOCUS AW361521 657 bp mRNA EST 04-FEB-2000
DEFINITION QV2-CT0261-261099-011-f03 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW361521
VERSION AW361521.1 GI:6866275
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 657)
HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1-QV2&l2-QV2-CT0261-
261099-011-f03&l3-1999-10-26&l4-1)
Seq primer: puc 18 forward
High quality sequence stop: 123
High quality sequence stop: 656.

FEATURES

source 1. 657
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site.1: Sma1; Site.2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 175 a 160 c 164 g 158 t
ORIGIN

Query Match 97.2%; Score 222.6; DB 10; Length 657;
Best Local Similarity 98.3%; Pred. No. 3.9e-58;
Matches 225; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GGCACAGTATGCTGGACAGACCGTGGGAAGACACTTTGTTTATCACCCTGGACA 60


```
Db 406 GGCACAGTGAATCGTGACACACCGCTGGGAAGACACTTGTCTTATCACCCTGGACA 465
OY 61 ACAGACCTCCCAAAATCTTCTGTGGATCCCGAGACAGAAAGAGGCTTGTGA 120
Db 466 ATGCAGCTCCCAAAATCTTCTGTGGATCCCGAGACAGAAAGAGGCTTGTGA 525
OY 121 GTGACAAAAACACCAAAATGGCTTACCTCAATCCAGCATTTGTAAGTTGGCACT 180
Db 526 GTGACAAAAACACCAAAATGGCTTACCTCAATCCAGCATTTGTAAGTTGGCACT 585
OY 181 TGGAAATACAGTCTGCAAGCAAGCTCAAAAACCTTGACCTGACTGTCA 229
Db 586 TGGAAATACAGTCTGCAAGCAAGCTCAAAAACCTTGACCTGACTGTCA 634

RESULT 8
AM361528/c 356 bp mRNA EST 04-FEB-2000
LOCUS AM361528 QV2-CT0261-261099-011-c04 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361528
VERSION AM361528.1 GI:6866282
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE HCGP http://www.ludwig.org.br/ORESTES.
JOURNAL 1 (bases 1 to 356)
COMMENT The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1-QV2&t2-QV2-CT0261-
261099-011-c04&t3-1999-10-26&t4-1)
Seq primer: puc 18 forward
High quality sequence start: 3
Location/Qualifiers
FEATURES
Source
1..356
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 79 a 73 c 107 g 97 t
ORIGIN
Query Match 87.7%; Score 200.8; DB 10; Length 356;
Best local Similarity 99.0%; Pred. No. 2e-51;
Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Db 236 GGCATCCAGTGGACAGACAGGAGCTTTGTAGTGACAAAAACCAAAATGGCCT 237
OY 346 ACCTCAAAATCCAGGATTTAGTGTGACCTTGGAAATACAGTCTGACCAAGCT 205
Db 236 ACCTCAAAATCCAGGATTTAGTGTGACCTTGGAAATACAGTCTGACCAAGCT 177
OY 236 CACAACCTTGACCTGACTGTCA 229
Db 176 CACAACCTTGACCTGACTGTCA 153

RESULT 9
AM361522 654 bp mRNA EST 04-FEB-2000
LOCUS AM361522 QV2-CT0261-261099-011-f10 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361522
VERSION AM361522.1 GI:6866276
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE HCGP http://www.ludwig.org.br/ORESTES.
JOURNAL 1 (bases 1 to 654)
COMMENT The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1-QV2&t2-QV2-CT0261-
261099-011-f10&t3-1999-10-26&t4-1)
Seq primer: puc 18 forward
High quality sequence start: 132
Location/Qualifiers
FEATURES
Source
1..654
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 173 a 156 c 162 g 163 t
ORIGIN
Query Match 74.5%; Score 170.6; DB 10; Length 654;
Best local Similarity 94.7%; Pred. No. 4.1e-42;
Matches 198; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
```


DB	142	ACCGATCCTCTCTACAAATATTATTATCTGGGATGCCAGCGGAGTGGAAACAAATGCTTTTATA	241
OY	191	GTGGACAAACAAACACCAAAATGCGCTACTCTCCAAATCCAGGACATCTGTAGTGGGACT	180
DB	242	CTAGACACACACCACTAAGGTGGCTTACTCTCCAAATGCGGACGCGGTAGTGGGCTTT	301
OY	191	TGGAAATACAGTCTGGCAAGCAAGCTCAACCAACTTGGACCTTGACATGCA	229
DB	302	TGGAAATACAGCATTCACAGCGAGCTCACACACTCTCACCTTGACATGCA	350
RESULT	11		
LOCUS	AK007465		
DEFINITION	AK007466	2915 bp	mRNA
ACCESSION	AK007466		
VERSION	AK007466.1		
KEYWORDS	GI:12841032		
SOURCE	AK007466.1		
ORGANISM	Mus musculus		
REFERENCE	AK007466	2915 bp	mRNA
AUTHORS	Mus musculus	10 day old male pancreas cDNA, RIKEN full-length	
JOURNAL	enriched library, clone:1810012P03, full insert sequence.		
MEDLINE	AK007466.1	GI:12841032	
PUBMED	99279253		
REFERENCE	10349636		
AUTHORS	(bases 1 to 2915)		
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
JOURNAL	Ittoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
MEDLINE	Normalization and subtraction of cap-trapper-selected cDNAs to		
PUBMED	prepare full-length cDNA libraries for rapid discovery of new genes		
REFERENCE	Genome research. 10 (10), 1617-1630 (2000)		
AUTHORS	11042159		
TITLE	(bases 1 to 2915)		
JOURNAL	Shibata, K., Ittoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
MEDLINE	Konno, H., Akiyama, K., Nishi, K., Kitsuana, T., Teshiro, H., Ittoh, M.,		
PUBMED	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
REFERENCE	Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
AUTHORS	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,		
TITLE	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,		
JOURNAL	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
MEDLINE	RIKEN integrated sequence analysis (RISA) system-384-format		
PUBMED	sequencing pipeline with 384 multiplexed sequencer		
REFERENCE	Genome research. 10 (11), 1757-1771 (2000)		
AUTHORS	20530913		
TITLE	Genome research. 10 (11), 1757-1771 (2000)		
JOURNAL	Genome research. 10 (11), 1757-1771 (2000)		
MEDLINE	Genome research. 10 (11), 1757-1771 (2000)		
PUBMED	Genome research. 10 (11), 1757-1771 (2000)		
REFERENCE	Genome research. 10 (11), 1757-1771 (2000)		
AUTHORS	Genome research. 10 (11), 1757-1771 (2000)		
TITLE	Genome research. 10 (11), 1757-1771 (2000)		
JOURNAL	Genome research. 10 (11), 1757-1771 (2000)		
MEDLINE	Genome research. 10 (11), 1757-1771 (2000)		
PUBMED	Genome research. 10 (11), 1757-1771 (2000)		
REFERENCE	Genome research. 10 (11), 1757-1771 (2000)		
AUTHORS	Genome research. 10 (11), 1757-1771 (2000)		
TITLE	Genome research. 10 (11), 1757-1771 (2000)		
JOURNAL	Genome research. 10 (11), 1757-1771 (2000)		
MEDLINE	Genome research. 10 (11), 1757-1771 (2000)		
PUBMED	Genome research. 10 (11), 1757-1771 (2000)		
REFERENCE	Genome research. 10 (11), 1757-1771 (2000)		
AUTHORS	Genome research. 10 (11), 1757-1771 (2000)		
TITLE	Genome research. 10 (11), 1757-1771 (2000)		
JOURNAL	Genome research. 10 (11), 1757-1771 (2000)		
MEDLINE	Genome research. 10 (11), 1757-1771 (2000)		
PUBMED	Genome research. 10 (11), 1757-1771 (2000)		
REFERENCE	Genome research. 10 (11), 1757-1771 (2000)		
AUTHORS	Genome research. 10 (11), 1757-1771 (2000)		
TITLE	Genome research. 10 (11), 1757-1771 (2000)		
JOURNAL	Genome research. 10 (11), 1757-1771 (2000)		
MEDLINE	Genome research. 10 (11), 1757-1771 (2000)		
PUBMED	Genome research. 10 (11), 1757-1771 (2000)		
REFERENCE	Genome research. 10 (11), 1757-1771 (2000)		
AUTHORS	Genome research. 10 (11), 1757-1771 (2000)		
TITLE	Genome research. 10 (11), 1757-1771 (2000)		
JOURNAL	Genome research. 10 (11), 1757-1771 (2000)		
MEDLINE	Genome research. 10 (11), 1757-1771 (2000)		
PUBMED	Genome research. 10 (11), 1757-1771 (2000)		
REFERENCE	Genome research. 10 (11), 1757-1771 (2000)		
AUTHORS	Genome research. 10 (11), 1757-1771 (2000)		
TITLE	Genome research. 10 (11), 1757-1771 (2000)		
JOURNAL			

REFERENCE	
AUTHORS	
TITLE	
1 (bases 1 to 2933)	
Carninci, P. and Hayashizaki, Y.	
High-efficiency full-length cDNA cloning	

source 1. .935
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="4215598"
/clone_lib="NCL_CGAP_CO24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: PCWV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCL_CGAP Library."

BASE COUNT 279 a 235 c 237 g 184 t
ORIGIN

Query Match 66.1%; Score 151.4; DB 11; Length 935;
Best Local Similarity 81.7%; Pred. No 3.5e-36;
Matches 187; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 1 GGCACAGTATCGTGGACAGCAGCGTGGAAAGACACTTTGTTCTTATCAGCTGACA 60
DB 79 GGCACAGTATCGTGGACAGCAGCGTGGAAAGACACTTTGTTCTTATCAGCTGACA 138
QY 61 AGCAGAGCTCCCAATCTCTCTGTGGATCCAGTGGACAGAAAGAGTGGCTTGTGA 120
DB 139 AGCAGAGCTCCCAATCTCTCTGTGGATCCAGTGGACAGAAAGAGTGGCTTGTGA 198
QY 121 GTGGACAAAACCAAAATGCTTACCTCAAAATCCAGGCAATGCTTAAGTTGGCACT 180
DB 199 CTAGACAAACCACTAAGGTGGCTTACCTCAAAATCCAGGCAATGCTTAAGTTGGC-TT 257
QY 181 TGGAAATACAGCTGCAAGCAAGCTCAAAACCTTGAACCTGACTGTCA 229
DB 258 TGGAAATACAGCAATTCAGGAGCTCAAGCTCTCACTGACTGTCA 306

RESULT 15
AW753448/c 496 bp mRNA EST 28-APR-2000
LOCUS AW753448
DEFINITION QV2-CT0261-261099-011-d04 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW753448
VERSION AW753448.1 GI:7668380
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 496)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-CT0261-
261099-011-d04&t3=1999-10-26&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 59
High quality sequence stop: 89.

TITLE
JOURNAL
MEDLINE
COMMENT

FEATURES
Location/Qualifiers
1. .496
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORSTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 104 a 116 c 128 g 148 t
ORIGIN

Query Match 60.7%; Score 139; DB 10; Length 496;
Best Local Similarity 99.3%; Pred. No 2.2e-32;
Matches 150; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 79 CTCTCTGGATCCAGTGGACAGAGAGTGGCTTTGTAGTGGACAAAACACCAAA 138
DB 496 CTCTCTGGATCCAGTGGACAGAGAGTGGCTTTGTAGTGGACAAAACACCAAA 438
QY 159 ATGGCTTACCTCCAAATCCAGGCAATGCTTAAGTTGGCACTTGAATACAGTGTCAA 198
DB 437 ATGGCTTACCTCCAAATCCAGGCAATGCTTAAGTTGGCACTTGAATACAGTGTCAA 378
QY 199 GCAAGCTCACAACCTTGAACCTGACTGTCA 229
DB 377 GCAAGCTCACAACCTTGAACCTGACTGTCA 347

Search completed: April 3, 2002, 20:27:33
Job time: 40216 sec

This Page Blank (uspto)

[illegible]

RESULT 2
5187077-16/c
Patent No. 5187077
APPLICANT: GEARING, DAVID P.; GOUGH, NICHOLAS M.; HILTON,
DOUGLAS J.; KING, JULIE A.; MCCALFE, DONALD NICE, EDOUARD C.
NICOLA, NICOS A.; SIMPSON, RICHARD J.; WILLSON, TRACY A.
TITLE OF INVENTION: LEUKEMIA INHIBITORY FACTOR
NUMBER OF SEQUENCES: 41
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/294,514
FILING DATE: 09-DEC-1988
SEQ ID NO.:16
LENGTH: 1437
5187077-16

	Query Match	Best Local Similarity	Matches	Score 12.9%; 61.8%;	Pred. No.	DB 0.35;	Length 1437;	Gaps 0
					Conservative	Mismatches	Indels	
OY	59	CACGCAAGCCTCCCAAATTCCTGTGGGATCCAGTGCACGAAGAAGCAAGTGCGTTG	118					
Db	1098	CAACCAAGCCTTCACAAGTTCTGTGAGACACAGCAAAAGAACCATGATGATGACTCTG	1039					
OY	119	TATGTGCACAAAACAC	134					
Db	1038	TAGGGGCTTAGACAC	1023					

RESULT 3
5427925-14/c
Patent No. 5427925
APPLICANT: GEARING, DAVID P.; GOUGH, NICHOLAS M.; HILTON,
DOUGLAS T.; KING, JULIE A.; MERCALF, DONALD; NICE, EDOUARD C.
NICOLA, NICOS A.; SIMPSON, RICHARD J.; WILSON, TRACY A.
TITLE OF INVENTION: RECOMBINANT METHOD FOR MAKING
LEUKEMIA INHIBITOR FACTOR
NUMBER OF SEQUENCES: 38
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/58,979
FILING DATE: 06-MAY-1993
APPLICATION NUMBER: 1948,614
FILING DATE: 22-SEP-1992
APPLICATION NUMBER: 1667,159
FILING DATE: 11-MAR-1991
SEQ ID NO:14:
LENGTH: 1437
5427925-14

[illegible]

RESULT

US-09-188-930-28
Sequence 28, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murlison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
TITLE OF INVENTION: and Methods For Their Use
FILE: REFERENCE: 11000.1011c1
CURRENT FILING: DATE: 1998-11-09
CURRENT APPLICATION NUMBER: US/09/188,930A
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 28
LENGTH: 1123
TYPE: DNA
ORGANISM: Rat
US-09-188-930-28

Query Match	12.7%;	Score 29;	DB 3;	Length 1123;
Best Local Similarity	63.8%;	Pred. No. 0.51;	25;	Indels 0; Gaps
Matches 44;	Conservative	0;	Mismatches	0;
Q7	127	AAAMACCAAAATGGCCTACTCTCAAAATCCAGCATTTGTAAGTTGGCACTTGGAAA	186	
Db	757	aaacaagaggaataatcagatataccagcaaacatgcagtcgtgaattggccaaaggaga	826	
Q7	187	TACAGTCG	195	
Db	837	tgaagtcctg	835	

RESULT 5
 US-09-188-930-203
 : Sequence 203, Application US/09188930A
 : Patent No. 6150502
 : GENERAL INFORMATION:
 : APPLICANT: Watson, James D.
 : APPLICANT: Strachan, Lorna
 : APPLICANT: Sleeman, Matthew
 : APPLICANT: Ormstedt, Rene
 : APPLICANT: Murison, James Greg
 : TITLE OF INVENTION: Compositions Isolated From Skin Cells
 : FILE REFERENCE: 11000.1011c1
 : CURRENT APPLICATION NUMBER: US/09/188,930A
 : CURRENT FILING DATE: 1998-11-09
 : NUMBER OF SEQ ID NOS: 348
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 203
 : LENGTH: 1123
 : TYPE: DNA
 : ORGANISM: Rat
 US-09-188-930-203

	Query Match	12.7%	Score 29	DB 3	Length 1123
Best Local Similarity	63.8%	Pred. No. 0.51	25	Indels 0	Gaps
Matches 44	Conservative	0	Mismatches	0	Gaps
0Y	127	AAAAACACCAAAATGCGCTACCTCCAAATGCCAGCATTTGCTAAAGTTGGCACTTGGAAA	186		
DB	757	aacaaagggaaatcatagataatccagcaaaccaatgcgttcgaagtttggccaaagagga	826		
0Y	187	TACAGTCG	195		
DB	827	tgaagtcg	835		


```

US-09-116-498-5/C
Sequence 5, Application US/09116498
Patent No. 6251582
GENERAL INFORMATION:
APPLICANT: Littman, Dan R.
Deng, Hongkui
Yountmaz, Derya
Ramanji, Vinet N.K.
TITLE OF INVENTION: NOVEL ALTERNATIVE G-COUPLED RECEPTORS
ASSOCIATED WITH RETROVIRAL ENTRY INTO CELLS, METHODS OF
IDENTIFYING THE SAME, AND DIAGNOSTIC AND THERAPEUTIC USES
THEREOF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/116,498
FILING DATE: 16-Jul-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEX: 201-343-1684
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 base pairs
TYPE: nucleic acid
STRANDNESS: double
MOLECULE TYPE: cDNA
TOPOLOGY: linear
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: pigtail macaque
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-116-498-5

Query Match      11.8%; Score 27; DB 4; Length 1029;
Best Local Similarity 49.6%; Pred. No. 2.7;
Matches 69; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

OY   63 GCAGCGCTCCCAAAATCTTCGTGGGATCCCAGTCGACAGACAAGCAAGTGCGCTTTGTACT 122
     ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    613 GAGGTGGCAAGAAGAACCCAGTCATCTGTGGTGGCAAGAACCAACAGTAATCTCT 554

OY   123 GGACAAAACACCACCAAAATGGCTTACTCTCAAATCCAGGCATTTGCTAAGTTGGCACTTG 182
     || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    553 TGTCATGATTAACCAACATATGACGCTGTGCCAGATTAAAGACATTCGCATAGATAATTGGG 494

OY   183 GAATACAGCTCTGCAGCA 201
     | | | | | | | | | |
Db    493 GCAGGAACACGACGAGGA 475

```

GENERAL INFORMATION:
APPLICANT: Surli, Bruno
TITLE OF INVENTION: Bacterial Vectors
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA

SIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/186,222
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/672,205
FILING DATE: 19-MAR-1991
APPLICATION NUMBER: GB 9006400.7
FILING DATE: 22-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Villamizar, Joann
REGISTRATION NUMBER: 30,598
REFERENCE/DOCKET NUMBER: 4-17994/A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914)785-7121
TELEFAX: (914)347-5769

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1920 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Lactococcus lactis LM0230
INDIVIDUAL ISOLATE: Major Secretion Product (MSP) Gene
IMMEDIATE SOURCE:
CLONE: pUCRS

FEATURES:
NAME/KEY: CDS
LOCATION: 411..1793
US-08-186-222-1

Query Match 11.6%; Score 26.6; DB 1; Length 1920;
Best Local Similarity 56.2%; Pred. No. 5.2;
Matches 50; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

D1 138 AAAACCAAAATGGCCCTACCTCCAAATCCAGGCGATTCGTAAAGTTGGCATTTGGAAT 187
|||||
D2 611 AAATGCCATAAATCAAAGCCGAACCTAAAGCGACTTAATGCCTCAAAATGCTACTTGAACG 690
|||||
OY 138 ACAGTCTCAAGCAAGCTCACAAACCTTG 216
|||||
D3 681 AAGTATCAAGAAGACGTACAAGAAGACTTG 719
|||||

RESULT 8
US-07-632-292-1/C
Sequence 1, Application US/07882292
Patent No. 5324638
GENERAL INFORMATION:
APPLICANT: Tao, Wufan
APPLICANT: Lai, Eseng
TITLE OF INVENTION: BRAIN TRANSCRIPTION FACTOR, NUCLEIC ACIDS
TITLE OF INVENTION: ENCODING SAME AND USES THEREOF

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: c/o Cooper and Dunham, 30 Rockefeller
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/882,292
FILING DATE: 19920513
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41472
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-664-0525
TELEFAX: 212-664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2830 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: CDS
LOCATION: 443..1882
OTHER INFORMATION:
FEATURE:
NAME/KEY: misc.feature
LOCATION: 926..1255
OTHER INFORMATION: /note= "nucleotide sequence encoding DNA
OTHER INFORMATION: binding domain homology"
FEATURE:
NAME/KEY: misc.signal
LOCATION: 1883..1885
OTHER INFORMATION: /note= "translation termination codon"
US-07-882-292-1

Query Match 11.5%; Score 26.4; DB 1; Length 2830;
Best Local Similarity 55.4%; Pred. No. 7.5;
Matches 51: Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 101 AGACGAAGTGGCTTTGAGTGGACAAACACCAAAATGGCTTACCTCCAAATCCGAG 160
DB 2539 AAAAATAAGTACCTGTTAGTGGACCACTACATCGAATATATACACAACTGAAG 2540

QY 161 GCATTGCTAAGTGGCACTTGGAAATACAGT 192
DB 2539 GCAATCCTTAATTTGTCTTTCGGATTCAAT 2508

RESULT 9
US-08-331-644-1/c
Sequence 1, Application US/08331644
Patent No. 5976872
GENERAL INFORMATION:
APPLICANT: Tao, Wufan
APPLICANT: Lai, Eseng
TITLE OF INVENTION: BRAIN TRANSCRIPTION FACTOR, NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING SAME AND USES THEREOF

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,644
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/882,292
FILING DATE: 13-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41472-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2830 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: CDS
LOCATION: 443..1882
OTHER INFORMATION:
FEATURE:
NAME/KEY: misc.feature
LOCATION: 926..1255
OTHER INFORMATION: /note= "nucleotide sequence
OTHER INFORMATION: encoding DNA binding domain
OTHER INFORMATION: homology"
FEATURE:
NAME/KEY: misc.signal
LOCATION: 1883..1885
OTHER INFORMATION: /note= "translation termination
OTHER INFORMATION: codon"
US-08-331-644-1

Query Match 11.5%; Score 26.4; DB 2; Length 2830;
Best Local Similarity 55.4%; Pred. No. 7.5;
Matches 51: Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 101 AGACGAAGTGGCTTTGAGTGGACAAACACCAAAATGGCTTACCTCCAAATCCGAG 160
DB 2539 AAAAATAAGTACCTGTTAGTGGACCACTACATCGAATATATACACAACTGAAG 2540

QY 161 GCATTGCTAAGTGGCACTTGGAAATACAGT 192
DB 2539 GCAATCCTTAATTTGTCTTTCGGATTCAAT 2508

RESULT 10
PCT-US93-04102-1/c
Sequence 1, Application PC/TUS9304102
GENERAL INFORMATION:
APPLICANT: Tao, Wufan
APPLICANT: Lai, Eseng

TITLE OF INVENTION: BRAIN TRANSCRIPTION FACTOR, NUCLEIC
ACIDS ENCODING SAME AND USES THEREOF
FILE REFERENCE: Strawberry Fruit Promoters for Gene Expression
CURRENT APPLICATION NUMBER: US/09/020,033A
CURRENT FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 1
LENGTH: 2061
TYPE: DNA
ORGANISM: Fragaria vesca
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(2061)
FEATURE:
NAME/KEY: unsure
LOCATION: (897)
OTHER INFORMATION: position 897 can be represented by any nucleotide
US-09-020-033-1
CLASSIFICATION:
FILING DATE: 19930430
APPLICATION NUMBER: PCT/US93/04102
CURRENT APPLICATION DATA:
SOFTWARE: Patentln Release #1.0, Version #1.25
OPERATING SYSTEM: PC-DOS/MS-DOS
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
COUNTRY: USA
ZIP: 10112
STATE: New York
CITY: New York
STREET: c/o Cooper and Dunham, 30 Rockefeller Plaza
ADDRESS: John P. White
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 4
TITLE OF INVENTION: ACIDS ENCODING SAME AND USES THEREOF
PCT-US93-04102-1
OTHER INFORMATION: /note="translation termination codon"
NAME/KEY: CDS
LOCATION: 443..1882
OTHER INFORMATION:
FEATURE:
NAME/KEY: misc-feature
LOCATION: 926..1255
OTHER INFORMATION: /note="nucleotide sequence encoding DNA
OTHER INFORMATION: binding domain homology"
FEATURE:
NAME/KEY: misc-signal
LOCATION: 1883..1885
OTHER INFORMATION:
PCT-US93-04102-1
Query Match 11.5%; Score 26.4; DB 5; Length 2830;
Best Local Similarity 55.4%; Pred. No. 7.5;
Matches 51; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 101 AGAAGAGGTGGCTTGTGTGTGACAAACCAACCAATGGCTTACTCCAAATCCAG 160
DB 2599 AAAAATTAAGTGAACCTGTGTGTGACCAATGCAATGCAAAATATACCAACAACTGAG 2540
QY 161 GCATTGCTAAGTTGGCACTTGGCAATACAGT 192
DB 2539 GCAATCTTAATTTTGTCTTTCGATTCAT 2508
RESULT 11
US-09-020-033-1/c
Sequence 1, Application US/09020033A
Patent No. 6043410
GENERAL INFORMATION:

APPLICANT: Wilkinson, Jack O.
TITLE OF INVENTION: Strawberry Fruit Promoters for Gene Expression
FILE REFERENCE: Strawberry Promoters for Gene Expression
CURRENT APPLICATION NUMBER: US/09/020,033A
CURRENT FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 1
LENGTH: 2061
TYPE: DNA
ORGANISM: Fragaria vesca
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(2061)
FEATURE:
NAME/KEY: unsure
LOCATION: (897)
OTHER INFORMATION: position 897 can be represented by any nucleotide
US-09-020-033-1
CLASSIFICATION:
FILING DATE: 19930430
APPLICATION NUMBER: PCT/US93/04102
CURRENT APPLICATION DATA:
SOFTWARE: Patentln Release #1.0, Version #1.25
OPERATING SYSTEM: PC-DOS/MS-DOS
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
COUNTRY: USA
ZIP: 10112
STATE: New York
CITY: New York
STREET: c/o Cooper and Dunham, 30 Rockefeller Plaza
ADDRESS: John P. White
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 4
TITLE OF INVENTION: ACIDS ENCODING SAME AND USES THEREOF
PCT-US93-04102-1
OTHER INFORMATION: /note="translation termination codon"
NAME/KEY: CDS
LOCATION: 443..1882
OTHER INFORMATION:
FEATURE:
NAME/KEY: misc-feature
LOCATION: 926..1255
OTHER INFORMATION: /note="nucleotide sequence encoding DNA
OTHER INFORMATION: binding domain homology"
FEATURE:
NAME/KEY: misc-signal
LOCATION: 1883..1885
OTHER INFORMATION:
PCT-US93-04102-1
Query Match 11.4%; Score 26.2; DB 3; Length 2061;
Best Local Similarity 54.7%; Pred. No. 7.5;
Matches 52; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 138 CTTTGTCTTATCACCCTGGACACGAGCCCTCCCAATCTTCTGGATCCAGTG 97
DB 903 CATCTCTCTATCCTGCTTCAACGCTCAAGGTTCCATCTCAGGAGCTG 934
QY 98 GACAGAACGAGTGGCTTGTGTGTGACAAAC 132
DB 933 GACATATCGAAGCAGATTCCAGTGAGAAATC 899
RESULT 112
US-09-027-381-1
Sequence 1, Application US/09027381
Patent No. 6075137
GENERAL INFORMATION:
APPLICANT: CULP, JEFFREY
APPLICANT: MCNUITY, DEAN
APPLICANT: ELLIS, CATHERINE
TITLE OF INVENTION: HUMAN UROTENSIN II
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/027,381
FILING DATE: 20-FEB-1998
CLASSIFICATION:
PCT-US93-027-381-1
OTHER INFORMATION:
NAME/KEY: promoter
LOCATION: (1)..(2061)
FEATURE:
NAME/KEY: unsure
LOCATION: (897)
OTHER INFORMATION: position 897 can be represented by any nucleotide
US-09-020-033-1
CLASSIFICATION:
FILING DATE: 19930430
APPLICATION NUMBER: PCT/US93/04102
CURRENT APPLICATION DATA:
SOFTWARE: Patentln Release #1.0, Version #1.25
OPERATING SYSTEM: PC-DOS/MS-DOS
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
COUNTRY: USA
ZIP: 10112
STATE: New York
CITY: New York
STREET: c/o Cooper and Dunham, 30 Rockefeller Plaza
ADDRESS: John P. White
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 4
TITLE OF INVENTION: ACIDS ENCODING SAME AND USES THEREOF
PCT-US93-04102-1
OTHER INFORMATION: /note="translation termination codon"
NAME/KEY: CDS
LOCATION: 443..1882
OTHER INFORMATION:
FEATURE:
NAME/KEY: misc-feature
LOCATION: 926..1255
OTHER INFORMATION: /note="nucleotide sequence encoding DNA
OTHER INFORMATION: binding domain homology"
FEATURE:
NAME/KEY: misc-signal
LOCATION: 1883..1885
OTHER INFORMATION:
PCT-US93-04102-1
Query Match 11.5%; Score 26.4; DB 5; Length 2830;
Best Local Similarity 55.4%; Pred. No. 7.5;
Matches 51; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 101 AGAAGAGGTGGCTTGTGTGTGACAAACCAACCAATGGCTTACTCCAAATCCAG 160
DB 2599 AAAAATTAAGTGAACCTGTGTGTGACCAATGCAATGCAAAATATACCAACAACTGAG 2540
QY 161 GCATTGCTAAGTTGGCACTTGGCAATACAGT 192
DB 2539 GCAATCTTAATTTTGTCTTTCGATTCAT 2508
RESULT 11
US-09-020-033-1/c
Sequence 1, Application US/09020033A
Patent No. 6043410
GENERAL INFORMATION:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 688 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-027-381-1

Query Match 11.4%; Score 26; DB 3; Length 688;
Best Local Similarity 65.5%; Pred. No. 5.1;
Matches 38; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 13 GTGGACAGCAGCGTGGAAAGAGACACTTTGTTTATCATCCTGGAACAGCAGCCTC 70
DB 621 GTGGAGAAAACACTAGCAACACTGTTGATCTTACTGGAATAATATCTCTC 678

RESULT 13
US-08-860-174A-3/C
Sequence 3, Application US/08860174A
Patent No. 5989830

GENERAL INFORMATION:

APPLICANT: DAVIS, Paul James
APPLICANT: VAN DER LOGT, Cornelis Paul Erik
APPLICANT: VERHOEIJEN, Martine Elisa
APPLICANT: WILSON, Steve
TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT
TITLE OF INVENTION: ANTIBODY FRAGMENT ANALOGUE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP
STREET: 1100 New York Avenue, N.W.
CITY: WASHINGTON, D.C.
STATE:
COUNTRY: UNITED STATES
ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS DOS Text
SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/860,174A
FILING DATE: June 16, 1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95307332.7
FILING DATE: October 16, 1995
APPLICATION NUMBER: PCT/EP/96/03605
FILING DATE: August 14, 1996

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "CDNA domains with synthetic
linker(s)"
IMMEDIATE SOURCE:
CLONE: HindIII-EcoRI insert Fv.3418
FEATURE:

NAME/KEY: CDS
LOCATION: 36..443
OTHER INFORMATION: /product= "pelB-VH3418"

FEATURE:
NAME/KEY: sig_peptide
LOCATION: 36..101

OTHER INFORMATION: /product= "pectate lyase"

FEATURE:
NAME/KEY: mat_peptide
LOCATION: 102..440
OTHER INFORMATION: /product= "VL4318"

FEATURE:
NAME/KEY: CDS
LOCATION: 495..884
OTHER INFORMATION: /product= "VL4318"

LOCATION: 102..440
OTHER INFORMATION: /product= "VH3418"

FEATURE:
NAME/KEY: CDS

LOCATION: 495..884

OTHER INFORMATION: /product= "pelB-VL4318"

FEATURE:
NAME/KEY: sig_peptide

LOCATION: 495..560

OTHER INFORMATION: /product= "pectate lyase"

Query Match 11.4%; Score 26; DB 2; Length 920;
Best Local Similarity 52.8%; Pred. No. 5.9;
Matches 56; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 139 GTCGCTTTGTAGTGGACAAAACCAAAATGGCTACCTCCAAATCCAGGCAATGCT 168
DB 543 GCTGGTGGCAGCAGATTAATACATCCAGCGCTGCCGTAGGCAATGCTATTCT 494

CY 169 AAGTTGGCACTTGGAAATACAGTCTGCAGCAGCAAGCTCCAAACT 214
DB 493 ATGACTGTCTCCTTGAATATGATTTGCATGCAAGCTCCACTACT 448

RESULT 14
US-09-171-025-24/C

Sequence 24, Application US/09171025
Patent No. 6239259

GENERAL INFORMATION:

APPLICANT: Multivalent and multispecific antigen-binding
TITLE OF INVENTION: protein
NUMBER OF SEQUENCES: 27
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,025
FILING DATE:
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:

LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "CDNA domains with synthetic linker(s)"
IMMEDIATE SOURCE:
CLONE: HindIII-EcoRI insert Fv.3418
FEATURE:

NAME/KEY: CDS
LOCATION: 36..443
OTHER INFORMATION: /product= "pelB-VH3418"

FEATURE:
NAME/KEY: sig_peptide
LOCATION: 36..101

OTHER INFORMATION: /product= "pectate lyase"

FEATURE:
NAME/KEY: mat_peptide
LOCATION: 102..440

OTHER INFORMATION: /product= "VH3418"

FEATURE:
NAME/KEY: CDS
LOCATION: 495..884
OTHER INFORMATION: /product= "VL4318"

FEATURE:
NAME/KEY: sig_peptide
LOCATION: 495..560
OTHER INFORMATION: /product= "pectate lyase"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 561..881
OTHER INFORMATION: /product= "VL3418"
US-09-171-025-24

Query Match 11.4%; Score 26; DB 4; Length 920;
Best Local Similarity 52.8%; Pred. No. 5.9;
Matches 56; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 109 GGTGCTTGTAGTGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCATTGCT 168
DB 553 GCTGCTTGGGACGACGATTAATACATCCAGCGCTGCCCTAGGCAATAGTATTTCATT 494
OY 169 AAGTTGGCACTTGAATAATACAGCTGCAAGCAAGCTCACAAACCT 214
DB 493 ATGACTGTCTCTTGAATAATGATTGCAATGCAAGCTCCCATAGCT 448

RESULT 15
US-08-894-922A-4/C
Sequence 4, Application US/08894922A
Patent No. 5863765
GENERAL INFORMATION:
APPLICANT: BERRY, Mark John
APPLICANT: DAVIS, Paul James
APPLICANT: VAN DER LOGT, Cornelius P.E.
APPLICANT: WHITTELAM, Gary Clark
TITLE OF INVENTION: PRODUCTION IN YEASTS OF STABLE ANTIBODY
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,922A
FILING DATE: 03-SEP-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9504344.4
FILING DATE: 03-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00468
FILING DATE: 01-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul K.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 60113/241261
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)-861-3503
TELEFAX: (202)-823-0944
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 996 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA
US-08-894-922A-4

FRAGME

Query Match 11.4%; Score 26; DB 2; Length 996;
Best Local Similarity 52.8%; Pred. No. 6.1;
Matches 56; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 109 GGTGCTTGTAGTGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCATTGCT 168
DB 553 GCTGCTTGGGACGACGATTAATACATCCAGCGCTGCCCTAGGCAATAGTATTTCATT 510
OY 169 AAGTTGGCACTTGAATAATACAGCTGCAAGCAAGCTCACAAACCT 214
DB 509 ATGACTGTCTCTTGAATAATGATTGCAATGCAAGCTCCCATAGCT 464

Search Completed: April 3, 2002, 20:54:01
Job time: 41799 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:17:53 ; Search time 1321.64 Seconds
(without alignments)
148.548 Million cell updates/sec

Title: US-09-049-696-10

Perfect score: 229
Sequence: 1 GGCACACTGATCGTGACAG.....AACCTTGACCTGACTGTCA 229

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq.1101.*

1: /SIDSI/gcgdata/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	229	100.0	2742	22	AAH46102 Human C1CA1 coding
2	229	100.0	2745	20	AAH81927 Human ICACC-1 nucl
3	229	100.0	2825	22	AAH46124 Human C1CA1 cDNA
4	229	100.0	2854	22	AAH34879 Human colon cancer
5	229	100.0	2854	22	AAH81787 Human secreted pro
6	229	100.0	2867	22	AAH33285 Human colon cancer
7	229	100.0	3109	22	AAH35019 Human colon cancer
8	229	100.0	3111	20	AAZ09840 Human membrane spa
9	229	100.0	3311	22	AAH29502 C902P determined c
10	161.8	70.7	2739	22	AAH46101 Mouse Gob-5 coding
11	161.8	70.7	2843	22	AAH46120 Mouse Gob-5 cDNA

12	161.8	70.7	2931	20	AAH81925 Murine ICACC-1 nuc
13	107.4	46.9	2616	21	AAH64335 Clone 2516888 of a
14	107.4	46.9	3265	21	AAZ65095 Membrane-bound pro
15	107.4	46.9	2092	22	AAH92092 Human PRO1124 cDNA
16	107.4	46.9	3265	22	AAH44241 Human PRO1124 (UNQ
17	106.6	46.6	1802	21	AAH98067 Human colon cancer
18	106.6	46.6	1802	22	AAH33192 Human colon cancer
19	88	38.4	486	20	AAH46121 Mouse Gob-5 cDNA h
20	34.2	14.9	585	20	AAZ24629 Human lung cancer-
21	34.2	14.9	590	21	AAH65868 Human lung cancer-
22	34.2	14.9	2454	22	AAH82881 Human tumor-associ
23	34.2	14.9	2784	20	AAZ24658 Human lung tumor a
24	34.2	14.9	2784	21	AAH65897 Human lung cancer-
25	34.2	14.9	3156	21	AAZ51625 Human lung cancer-
26	34.2	14.9	3362	21	AAZ24657 Human lung tumor a
27	34.2	14.9	3362	21	AAH65896 Human lung cancer-
28	34.2	14.9	3351	20	AAZ24653 Human lung tumor a
29	34.2	14.9	3951	21	AAH65892 Human lung cancer-
30	34.2	14.9	8031	21	AAH65950 Human lung cancer-
31	33	14.4	3190	20	AAH81926 Human ICACC-2 nucl
32	31.6	13.8	4771	18	AAV02186 Human type C lecti
33	31.6	13.8	4771	21	AAH92447 Human type C lecti
34	31.6	13.8	5747	21	AAH77360 Human ORF2915
35	30.6	13.4	755	22	AAH98782 Human ORF2915
36	29.6	12.9	547	20	AAZ33810 Human EST-derived
37	29.6	12.9	138	9	AAH81177 Tobacco plant resi
38	29.6	12.9	3694	21	AAH64282 Insert from PLIFNK
39	29.6	12.9	4837	22	AAH59280 Human membrane-ass
40	29.6	12.9	4956	22	AAH61066 Human polynucleoti
41	29.4	12.8	300	20	AAZ13368 Human poly nucleoti
42	29.4	12.8	1196	21	AAH27335 Human gene express
43	29.4	12.8	1436	21	AAH27334 Rice RAB anther-sp
44	29.4	12.8	3003	21	AAH27333 Rice RAB anther-sp
45	29	12.7	1035	22	AAH99776 Skin cell cDNA, SE

ALIGNMENTS

RESULT: 1

ID AAH46102 standard; DNA; 2742 BP.

XX AAH46102;

DT 11SEP-2001 (first entry)

XX Human C1CA1 coding sequence, Seq ID NO:4.

DE Human C1CA1; goblet cell; mouse Gob-5 orthologue; drug screening;

KW expression inhibition; antisense therapy; gene therapy;

KW chronic obstructive pulmonary disease; bronchial asthma; antitastmatic;

KW ds.

OS Homo sapiens.

XX Key

FT Key

FT Key

FT Key

FT Key

FT Key

FT Key

FT Key

FT Key

FT Key

Location/Qualifiers
1.2742
/*tag= a
/product- "Human C1CA1"
/note= "No stop codon given in the specification"

MO200138530-A1.

21-MAY-2001.

22-NOV-2000: 2000MO-JP08232.

24-NOV-1999: 99GP-0333479.

27-APR-2000: 2000GP-0127589.

(TAKE) TAKEDA CHEM IND LTD.

XX Nakanishi A, Morita S;
XX
XX MPI: 2001-355935/37.
DR P-PSDB: AAB73716.
XX
XX New antisense nucleotide, useful for treatment and prevention of
PT bronchial asthma and chronic obstructive pulmonary disease
XX
XX
PS Claim 3; Page 82-84; 104pp; Japanese.
XX
XX The invention relates to an antisense nucleotide targeted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the Gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
CC pulmonary disease. The present sequence represents the human CLCA1
CC gene coding sequence.
CC
XX
SQ Sequence 2742 BP; 833 A; 616 C; 623 G; 670 T; 0 other;

Query Match 100.0%; Score 229; DB 22; Length 2742;
Best Local Similarity 100.0%; Pred. No. 3.4e-69;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCACAGTATCGTGGACAGACCGTGGGAAAGACACTTTGTTTATCACTGGACA 60
DB 1510 ggcacagtatcggtggacagaccgctgggaaagacacttgttctatcacctggaca 1569
OY 61 AGCAGACCTCCCAAAATCTTCTCTGGATCCATGGACAGACAGAGTGGCTTGA 120
DB 1570 agcagacctcccaaaatcttctctggatcccatggacagagtggtcttga 1629
OY 121 GTGGACAAAACACCAAAATGGCTACCTCCAAATCCAGGCTTGTAAAGTTGGCACT 180
DB 1630 gtggacaaaacaccaaattggctacctccaaatccaggcttgttaagtggcaact 1689
OY 181 TGGAAATACAGTCTGCAGACAGCTCACAACCTTGACCTGACTGTCA 229
DB 1690 tggaaatacagctgcagacagctcacaaccttgacctgactgtca 1738

RESULT 2
AAH81927
ID AAH81927 standard; cDNA; 2745 BP.
XX
XX
AC AAH81927;
XX
XX 13-JUN-2001 (first entry)
XX
XX Human ICACC-1 nucleotide sequence.
XX
XX ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
KW interleukin 9 induced calcium activated channel; IL-9;
KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
KW anti-inflammatory; immunomodulatory; cystic fibrosis;
KW inflammatory bowel disease; autoimmune disease; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FH 1..2745
XX CDS /*tag= a
XX /product= "ICACC-1"
XX /note= "IL-9 induced calcium activated chloride channel"
XX
XX W09944620-A1.

PD 10-SEP-1999.
XX
XX 03-MAR-1999; 99MO-US04703.
XX
XX 03-MAR-1998; 98US-0076815.
XX
XX (MAGA-) MAGANIN PHARM INC.
XX
XX
PI Holroyd KJ, Alevitt RC, Maloy WL, Louhed J, McLane M;
PI Nicolaides NC, Zhou Y, Dong Q;
XX
XX MPI: 1999-550979/46.
DR P-PSDB: AAB74824.
XX
XX New nucleic acid encoding calcium activated chloride channel, used to
PT identify, e.g. specific modulators for treating atopic allergy -
XX
XX
PS Claim 1; Fig 4B; 75pp; English.
XX
XX The present sequence encodes the human interleukin 9 (IL-9) induced
CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
CC have anti-allergic, anti-asthmatic, anti-inflammatory and
CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
CC used to upregulate asthma (or more generally atopic allergy), while those
CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
CC Menstrual levels of ICACC-1 can be used in the diagnosis of asthma
CC (increased levels) or IBD (reduced levels), also for monitoring
CC treatment of these conditions. The ICACC proteins can be used:
CC (i) to raise specific antibodies (Ab), useful: (a) as immunoassay
CC reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
CC (A) to disrupt bonding between ICACC and its ligands, and (iii) to be
CC identify modulators and binding partners. ICACC polynucleotides can be
CC used to generate transgenic animals or recombinant cells, used to screen
CC for antagonists, also as a source of therapeutic antisense agents or
CC diagnostic probes (for quantifying mRNA expression, e.g. for
CC identification of modulators).
XX
SQ Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;

Query Match 100.0%; Score 229; DB 20; Length 2745;
Best Local Similarity 100.0%; Pred. No. 3.4e-69;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCACAGTATCGTGGACAGACCGTGGGAAAGACACTTTGTTTATCACTGGACA 60
DB 1510 ggcacagtatcggtggacagaccgctgggaaagacacttgttctatcacctggaca 1569
OY 61 AGCAGACCTCCCAAAATCTTCTCTGGATCCATGGACAGACAGAGTGGCTTGA 120
DB 1570 agcagacctcccaaaatcttctctggatcccatggacagagtggtcttga 1629
OY 121 GTGGACAAAACACCAAAATGGCTACCTCCAAATCCAGGCTTGTAAAGTTGGCACT 180
DB 1630 gtggacaaaacaccaaattggctacctccaaatccaggcttgttaagtggcaact 1689
OY 181 TGGAAATACAGTCTGCAGACAGCTCACAACCTTGACCTGACTGTCA 229
DB 1690 tggaaatacagctgcagacagctcacaaccttgacctgactgtca 1738

RESULT 3
AAH46124
ID AAH46124 standard; cDNA; 2825 BP.
XX
XX
AC AAH46124;
XX
XX 11-SEP-2001 (first entry)
XX
XX Human CLCA1 cDNA, SEQ ID NO:26.
XX
XX Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
KW

expression inhibition; antisense therapy; gene therapy;
 chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
 ss.
 Homo sapiens.
 Location/Qualifiers
 Key 23..2767
 CDS /tag= a
 /product= "Human CLCA1"
 /transcript= (pos:476..478, aa:lys)
 WO200138530-A1.
 31-MAY-2001.
 22-NOV-2000; 2000MO-JP08232.
 24-NOV-1999; 99JP-0333479.
 27-APR-2000; 2000JP-0127589.
 (TAKE) TAKEDA CHEM IND LTD.
 Nakanishi A, Morita S;
 WPI: 2001-355935/37.
 P-PSDB: AAB73716.
 New antisense nucleotide, useful for treatment and prevention of
 bronchial asthma and chronic obstructive pulmonary disease -
 Example 5; Page 92-94; 104pp; Japanese.
 The invention relates to an antisense nucleotide targeted to the mouse
 Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 the CLCA1 gene (coding sequence shown in AAH46102). The invention also
 relates to an antibody specific for the Gob-5 protein, medical and
 diagnostic compositions containing the antisense nucleotide or the
 antibody, and methods and kits for screening for compounds which inhibit
 the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 The antisense oligonucleotides and antibody are therefore useful for the
 treatment and prevention of bronchial asthma and chronic obstructive
 pulmonary disease. The present sequence represents human CLCA1 cDNA.
 Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;
 Query Match 100.0%; Score 229; DB 22; Length 2825;
 Best Local Similarity 100.0%; Pred. No. 3,4e-69;
 Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGCACAGTGTGCTGGACAGCAGCGTGGCAAGACACTTGTTCATTACACCTGGACA 60
 Db 1532 ggcacagtgatcgtgacacacacgcgcyggaagacacttgccttaccacttgaca 1591
 Oy 61 ACGCAGCTCCGCCAAATCTCTCTGTGGATCCAGTGGACAGCAAGAGTGGCTTGTGA 120
 Db 1592 acgcagctcccccgaatcctctctctggtatccagtgacagaagcaggttgcttgyta 1651
 Oy 121 GTGACAAAAAACAACCAAAATGGCTTACTTCAATCCAGGCAATTGCTAAGTTGGCACT 180
 Db 1652 gtgacaaaaaacacaaatggccttacttcaatcccaagcattgtaagttggcact 1711
 Oy 181 TGGAAATACAGTCTGCAAGCAGCTCAAAACCTTGACCTGACTGCA 229
 Db 1712 tggaaatacagctctgcaagcagctcaaaaccttgaccctgactgca 1760
 RESULT 4
 AAH34879
 ID AAH34879 standard; cDNA; 2854 BP.
 AC AAH34879;

XX 23-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 1; ss.
 XX
 OS Homo sapiens.
 PN WO200122920-A2.
 05-APR-2001.
 28-SEP-2000; 2000MO-US26524.
 23-SEP-1999; 99US-0157137.
 03-NOV-1999; 99US-0163280.
 (HYMA-) HUMAN GENOME SCI INC.
 Ruben SM, Barash SC, Birse CE, Rosen CA;
 WPI: 2001-235357/24.
 P-PSDB: AAG75474.
 Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 useful for preventing, diagnosing and/or treating colorectal cancers -
 Claim 1; Page 3462-3463; 9803pp; English.
 AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 cancer-associated nucleic acid molecules (N) and proteins (P), where
 the proteins are collectively known as colon cancer antigens. The colon
 cancer antigens have cytostatic activity and can be used in gene
 therapy and vaccine production. N and P may be used in the prevention,
 diagnosis and treatment of diseases associated with inappropriate P
 expression. For example, N and P may be used to treat disorders
 associated with decreased expression by rectifying mutations or deletions
 in a patient's genome that affect the activity of P by expressing
 inactive proteins or to supplement the patient's own production of P.
 Additionally, N may be used to produce the colon cancer-associated P,
 by inserting the nucleic acids into a host cell and culturing the cell
 to express the proteins. N and P can be used in the prevention, diagnosis
 and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 and AAB77789 represent sequences used in the exemplification of the
 present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC rising at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;
 Query Match 100.0%; Score 229; DB 22; Length 2854;
 Best Local Similarity 100.0%; Pred. No. 3.5e-69;
 Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGCACAGTGTGCTGGACAGCAGCGTGGCAAGACACTTGTTCATTACACCTGGACA 60
 Db 1544 ggcacagtgatcgtgacacacacgcgcyggaagacacttgccttaccacttgaca 1603
 Oy 61 ACGCAGCTCCGCCAAATCTCTCTGTGGATCCAGTGGACAGCAAGAGTGGCTTGTGA 120
 Db 1604 acgcagctcccccgaatcctctctctggtatccagtgacagaagcaggttgcttgyta 1663
 Oy 121 GTGACAAAAAACAACCAAAATGGCTTACTTCAATCCAGGCAATTGCTAAGTTGGCACT 180
 Db 1664 gtgacaaaaaacacaaatggccttacttcaatcccaagcattgtaagttggcact 1723
 Oy 181 TGGAAATACAGTCTGCAAGCAGCTCAAAACCTTGACCTGACTGCA 229
 Db 1724 tggaaatacagctctgcaagcagctcaaaaccttgaccctgactgca 1772

RESULT 5
AA81787 standard; cDNA: 2854 BP.
ID AAF81787
XX
AC AAF81787;
XX
XX 12-JUN-2001 (first entry)
DE Human secreted protein gene 1 SEQ ID NO:11.
XX
XX Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;
KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
KW antialzheimer; antiparkinsonian; antimicrobial; vulnery; gene therapy;
KW immune disorder; hyperproliferative disorder; cardiovascular disease;
KW cancer; angiogenic disorder; neurological disorder; infectious disease;
KW wound healing; regeneration; chemotaxis; chromosome 1; ss.
XX
XX Homo sapiens.
OS
XX WO200112775-A2.
PN
XX 22-FEB-2001.
PD
XX 16-AUG-2000; 2000WO-US22325.
PF
XX 17-AUG-1999; 99US-0149182.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;
PI Birse CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;
PI P-PSDB; AAB74733.
DR
XX WPI: 2001-147550/15.
XX
XX Nucleic acids encoding 25 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
XX
XX Claim 1: Page 441; 485pp; English.
PS
XX AAF81787 to AAF81817 encode the human secreted proteins given in AAB74733
CC to AAB74772. Human secreted proteins can have activities based on the
CC tissues and cells they are expressed in. Example of activities include:
CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;
CC antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;
CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
CC anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and
CC vulnery. Human secreted proteins can be used in gene therapy and
CC vaccine. Human secreted protein nucleotide sequences (NAMI) and proteins
CC (PEP) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. For example, NAMI
CC and PEP may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of proteins by expressing inactive proteins or
CC to supplement the patient's own production of polypeptides. Disorders that
CC may be prevented, diagnosed and/or treated include immune disorders,
CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
CC angiogenic disorders, neurological disorders, infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to
CC AAF81786 and AAB74732 represent sequences used in the exemplification of
CC the present invention.
XX
XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match 100.0%; Score 229; DB 22; Length 2854;
Best Local Similarity 100.0%; Pred. No. 3.5e-69;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

07 1 GGACAGTGTGTCGACAGCAGCGTGGAAAGACACTTTGTTCTTACACTGACACA 60
Db |||||||
1544 ggcacagtgatcgtggacagcaccgctggaaaggacacttggttctctacacctggaca 1603
07 61 ACGCAGCCTCCCAAAATCTTCTGGAATCCAGTGACAGAACAGAGTGGCTTTGTA 120
Db |||||||
1604 acgacgctccccaatccctctctcgatccagtgagacaagaagctggcttggta 1663
07 121 GTGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTAAGTTGCACT 180
Db |||||||
1654 gtggacaaaacaccaaataggcctaccctccaaatccagcagctgctaaagttggcact 1723
07 181 TGGAAATGACGTCTGCAAGCAGCTCACAAACCTTGAACCTGACTGTCA 229
Db |||||||
1724 tggaaatcagctgcagcagcagctcacaacacttgacccctgactgtca 1772
RESULT 6
AAH3285
ID AAH3285 standard; cDNA: 2867 BP.
XX
XX AC AAH3285;
XX
XX 03-SEP-2001 (first entry)
DE
XX Human colon cancer antigen encoding cDNA SEQ ID NO:341.
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 1; ss.
XX
XX Homo sapiens.
OS
XX WO200122920-A2.
PN
XX 05-APR-2001.
PD
XX 28-SEP-2000; 2000WO-US26524.
PF
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruken SM, Barash SC, Birse CE, Rosen CA;
PI WPI: 2001-235357/24.
DR
XX P-PSDB; AAG73854.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1: Page 2452-2453; 9803pp; English.
PS
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;

Query Match 100.0%; Score 229; DB 22; Length 2867;
Best Local Similarity 100.0%; Pred. No. 3.5e-69;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCACAGTATGCTGGACAGCACCCTGGGAAAGACACTTTGTTTCTTATCACCCTGGACA 60
DB 1548 ggcacagtgatcgtgacagcaccgtgggaaaggacacttgtttctatcacctggaca 1607
OY 61 ACGCAGCCTCCCAAAATCCTTCTCTGGATCCAGTGACAGAGAAGGCTGCTTTGTA 120
DB 1608 acgcagcctcccaaatccttctcttgagatccagtgagacagaagaagtgagcttga 1667
OY 121 GTGACAAAACACCAAAATGCTCTACCTCCAAATCCCGAGCTTCTTAAGTTGGCACT 180
DB 1668 gtgacaaaacacccaatgacctaccccaatccagagcatgtctaagttggcact 1727
OY 181 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGACCTGACTGTCA 229
DB 1728 tggaaatacagtcctgcaagcagctcacaaccttgaccctgactgtca 1776

RESULT 7
AAH35019 standard; cDNA; 3109 BP.

XX AAH35019;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:2101.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma, ss.

XX Homo sapiens.

PN W0200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Blase CE, Rosen CA;

DR WPI: 2001-235357/24.

XX P-PSDB: AAG75614.

PS Claim 1; Page 3587-3588; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell

CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention. N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match 100.0%; Score 229; DB 22; Length 3109;
Best Local Similarity 100.0%; Pred. No. 3.6e-69;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCACAGTATGCTGGACAGCACCCTGGGAAAGACACTTTGTTTCTTATCACCCTGGACA 60
DB 1347 ggcacagtgatcgtgacagcaccgtgggaaaggacacttgtttctatcacctggaca 1456
OY 61 ACGCAGCCTCCCAAAATCCTTCTCTGGATCCAGTGACAGAGAAGGCTGCTTTGTA 120
DB 1457 acgcagcctcccaaatccttctcttgagatccagtgagacagaagaagtgagcttga 1516
OY 181 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGACCTGACTGTCA 229
DB 1517 gtgacaaaacacccaatgacctaccccaatccagagcatgtctaagttggcact 1576
OY 181 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGACCTGACTGTCA 229
DB 1577 tggaaatacagtcctgcaagcagctcacaaccttgaccctgactgtca 1625

RESULT 8
AAZ09849 standard; cDNA; 3111 BP.

XX AAZ09849;

DT 26-NOV-1999 (first entry)

DE Human membrane spanning protein MSP-5 cDNA fragment 2.

XX Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;
KM neoplastic disorder; immunological disorder; reproductive disorder;

XX MSP-5; ds.

XX Homo sapiens.

PN W05946380-A2.

PD 16-SEP-1999.

PF 09-MAR-1999; 99WO-US05073.

XX 13-MAR-1998; 98US-0039064.

XX (INCY-) INCYTE PHARM INC.

PI Tang YT, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;

DR WPI: 1999-551409/46.

XX P-PSDB: AAY33298.

PS New human membrane spanning proteins used to, e.g. prevent and treat
CC neoplastic disorders -

XX Example 1; Page 80-81; 81pp; English.

CC This invention describes novel human membrane spanning proteins (MSPs),
CC and the polynucleotides encoding them. The products of the invention are
CC used to diagnose, prevent and treat neoplastic, immunological and
CC reproductive disorders. This sequence encodes a human membrane spanning

CC protein MSP-5 fragment.
 XX Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;
 SQ

Query Match 100.0%; Score 229; DB 20; Length 3111;
 Best Local Similarity 100.0%; Pred. No. 3.6e-69;
 Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCACAGTGTGATGTCGACAGCACCCTGGGAAAGACACTTGTTCCTTATCAGCTGACACA 60
 1543 ggcacagtgatgtgtaagacacgctgggaaagacacttgttcttatacaccctggaca 1602
 61 ACGCAGCCTCCCAAAATCCTTCTGCGATCCCACTGACAGACAGAGTGCTTTGTA 120
 1603 acgcagcctccccaatccttctctggtggtccagtgacagaagcaagtggtcttgta 1662
 121 GTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGCATGCTAAGTTGGCACT 180
 1663 gtggacaaaacaccaaattggccttaccctccaaatccacagcatgtcaaggttggcact 1722
 181 TGGAAATACAGTGTGACAGCAGCTCACAACCTTGACCTGACTGCA 229
 1723 tggaaatacagtgtgacagcagctcacaaccttgaccctgactgtca 1771

RESULT 9
 AA129502
 ID AA129502 standard; cDNA; 3311 BP.
 XX
 AC AA129502;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE C902p determined cDNA sequence.
 XX
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 KM Immunogenic; gene therapy; vaccine; colonic cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200149716-A2.
 XX
 PD 12-JUL-2001.
 XX
 PE 29-DEC-2000; 2000MO-US35596.
 XX
 PR 30-DEC-1999; 9905-0476296.
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
 PI King GE, Wang T, Jiang Y;
 XX
 DR WPI: 2001-441847/47.
 XX
 PS Claim 2: Page 425-426; 472pp; English.
 XX
 CC The present invention describes colon tumour associated proteins (I) and
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate colon tumour associated protein (TCAP)
 CC expression, such as colonic cancer. For example, (I) and (II) may be

CC used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of TCAPs by expressing inactive proteins or to supplement the
 CC patient's own production of them. Additionally, (II) may be used to
 CC produce the TCAP proteins, by inserting the nucleic acids into a host
 CC cell culturing the cell to express the protein. (II) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC polymerase chain reaction (PCR) and hybridisation assays to detect and
 CC quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. (I) may
 CC also be used as antigens in the production of antibodies against TCAPs
 CC and in assays to identify modulators of TCAP expression and activity.
 CC Anti-(I) antibodies and antagonists may also be used to down regulate
 CC TCAP expression and activity. The anti-(I) antibodies may also be used
 CC as diagnostic agents for detecting the presence of TCAPs in samples
 CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
 CC and AA24494 to AA24523 represent nucleotide and amino acid sequences
 CC given in the exemplification of the present invention.
 XX
 SQ Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;

Query Match 100.0%; Score 229; DB 22; Length 3311;
 Best Local Similarity 100.0%; Pred. No. 3.7e-69;
 Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGCACAGTGTGATGTCGACAGCACCCTGGGAAAGACACTTGTTCCTTATCAGCTGACACA 60
 1861 ggcacagtgatgtgtaagacacgctgggaaagacacttgttcttatacaccctggaca 1920
 71 ACGCAGCCTCCCAAAATCCTTCTGCGATCCCACTGACAGACAGAGTGCTTTGTA 120
 1971 acgcagcctccccaatccttctctggtggtccagtgacagaagcaagtggtcttgta 1980
 121 GTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGCATGCTAAGTTGGCACT 180
 1981 gtggacaaaacaccaaattggccttaccctccaaatccacagcatgtcaaggttggcact 2040
 181 TGGAAATACAGTGTGACAGCAGCTCACAACCTTGACCTGACTGCA 229
 2041 tggaaatacagtgtgacagcagctcacaaccttgaccctgactgtca 2089

RESULT 10
 AAH46101
 ID AAH46101 standard; DNA; 2739 BP.
 XX
 AC AAH46101;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Mouse Gob-5 coding sequence, SEQ ID NO:3.
 XX
 KW Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
 KM expression inhibition; antisense therapy; gene therapy; bronchial asthma;
 KW chronic obstructive pulmonary disease; antiasthmatic; ds.
 XX
 OS Mus sp.
 XX
 PI Key Location/Qualifiers
 PI CDS 1..2739
 FT 1..2739
 FT /*tag=a
 FT /partial
 FT /product="Mouse Gob-5"
 FT /note="No stop codon given in the specification"
 XX
 XX WO200138530-A1.
 XX
 XX 31-MAY-2001.
 XX
 XX 22-NOV-2000; 2000MO-JP08232.
 XX
 XX 24-NOV-1999; 99JP-0333479.


```

PR 27-APR-2000; 2000JP-0127589.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Nakanishi A, Morita S;
PI WPI: 2001-355935/37.
XX P-PSDB: AAB73715.
DR
XX New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease -
XX
XX Claim 3; Page 80-82; 104pp; Japanese.
XX
XX The invention relates to an antisense nucleotide targeted to the mouse
XX Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX relates to an antibody specific for the Gob-5 protein, medical and
XX diagnostic compositions containing the antisense nucleotide or the
XX antibody, and methods and kits for screening for compounds which inhibit
XX the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX The antisense oligonucleotides and antibody are therefore useful for the
XX treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents the mouse Gob-5
XX gene coding sequence.
XX
SQ Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other;

Query Match 70.7%; Score 161.8; DB 22; Length 2739;
Best Local Similarity 81.7%; Pred. No. 6.5e-46;
Matches 187; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

OY 1 GGCACAGTATCGTGACAGACACCGTGGAAGACACTTGTGTTTCACTTGACCTGGACA 60
DB 1513 ggtcagtgatcgtgacagcgcgtggtggaagacacctgttctatcacctggaca 1572
OY 61 ACCGACCTCCCCCAATCCTTCTCTGGATCCAGTGGACAGACAGAGTGCGCTTTGTA 120
DB 1573 acgcatccctcacaatattatctggtatccagcgaggtgaaacaatggtttata 1632
OY 121 GTGGACAAAACACCAAAATGGCTACCTCCCAAGGCTTAAGTTGGCACT 180
DB 1633 ctgagacacaacctaaggtggtcaccctccagtcacgagcgttaagttgctt 1692
OY 181 TGGAAATACAGTCTGCAAGACGCTCACAACCTTGACCTGACTGCA 229
DB 1693 tggaaatacagcatcctaagcgagctcacagactctcaacttgactgtca 1741

RESULT 11
AAH46120
ID AAH46120 standard; cDNA; 2843 BP.
XX
XX AAH46120;
AC
XX 11-SEP-2001 (first entry)
DT
XX
XX Mouse Gob-5 cDNA, SEQ ID NO:22.
DE
XX
XX Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
KM expression inhibition; antisense therapy; gene therapy; bronchial asthma;
KM chronic obstructive pulmonary disease; antiasthmatic; ss.
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
PH CDS 15..2736
FT /*tag= a
FT /product= "Mouse Gob-5"
XX
XX WO200138530-A1.
XX

```

```

PD 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-JP08232.
XX
XX 24-NOV-1999; 99JP-0333479.
PR 27-APR-2000; 2000JP-0127589.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Nakanishi A, Morita S;
PI WPI: 2001-355935/37.
XX P-PSDB: AAB73715.
DR
XX New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease -
XX
XX Example 1; Page 89-91; 104pp; Japanese.
XX
XX The invention relates to an antisense nucleotide targeted to the mouse
XX Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX relates to an antibody specific for the Gob-5 protein, medical and
XX diagnostic compositions containing the antisense nucleotide or the
XX antibody, and methods and kits for screening for compounds which inhibit
XX the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX The antisense oligonucleotides and antibody are therefore useful for the
XX treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
XX
SQ Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;

Query Match 70.7%; Score 161.8; DB 22; Length 2843;
Best Local Similarity 81.7%; Pred. No. 6.6e-46;
Matches 187; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

OY 1 GGCACAGTATCGTGACAGACACCGTGGAAGACACTTGTGTTTCACTTGACCTGGACA 60
DB 1513 ggtcagtgatcgtgacagcgcgtggtggaagacacctgttctatcacctggaca 1586
OY 61 ACCGACCTCCCCCAATCCTTCTCTGGATCCAGTGGACAGACAGAGTGCGCTTTGTA 120
DB 1573 acgcatccctcacaatattatctggtatccagcgaggtgaaacaatggtttata 1646
OY 121 GTGGACAAAACACCAAAATGGCTACCTCCCAAGGCTTAAGTTGGCACT 180
DB 1617 ctgagacacaacctaaggtggtcaccctccagtcacgagcgttaagttgctt 1706
OY 181 TGGAAATACAGTCTGCAAGACGCTCACAACCTTGACCTGACTGCA 229
DB 1707 tggaaatacagcatcctaagcgagctcacagactctcaacttgactgtca 1755

RESULT 12
AAF81925
ID AAF81925 standard; cDNA; 2931 BP.
XX
XX AAF81925;
AC
XX 13-JUN-2001 (first entry)
DT
XX
XX Nucleic ICACC-1 nucleotide sequence.
DE
XX
XX ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
KM interleukin 9 induced calcium activated chloride channel; IL-9;
KM calcium activated chloride channel; anti-allergic; anti-asthmatic;
KM anti-inflammatory; immunomodulatory; cystic fibrosis;
KM inflammatory bowel disease; autoimmune disease; ss.
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
PH CDS
XX

```


RESULT 14
AA65095
ID AA65095 standard; cDNA; 3265 BP.
XX
AC AA65095;
XX
DT 05-APR-2000 (first entry)
XX
DE Membrane-bound protein PRO1124 encoding cDNA.
XX
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping; ss.
XX
OS Homo sapiens.
XX
PN MO963088-A2.
XX
PD 09-DEC-1999.
XX
PF 02-JUN-1999; 99MO-US12252.
XX
PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 05-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.

PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 25-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 17-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 18-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.

XX	(GETH) GENENTECH INC.
XX	Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PI	Wood WI, Yuan J;
XX	WPI: 2000-072883/06.
DR	P-PSDB: AAY66749.
XX	Membrane-bound proteins and related nucleotide sequences
PT	Claim 2; Fig 273; 822pp; English.
XX	The invention provides membrane-bound PRO polypeptides and
CC	polynucleotides encoding them. The PRO sequences of the invention were
CC	identified based on extracellular domain homology screening. The PRO
CC	sequences have homology with proteins including LDL receptors, TIE
CC	ligands and various enzymes. The membrane-bound proteins and receptor
CC	molecules are useful as pharmaceutical and diagnostic agents. Receptor
CC	immunoadhesins, for instance, can be used as therapeutic agents to block
CC	receptor-ligand interactions. The membrane-bound proteins can also be
CC	employed for screening of potential peptide or small molecule inhibitors
CC	of the relevant receptor/ligand interaction. The PRO encoding sequences
CC	are useful as hybridization probes, in chromosome and gene mapping and in
CC	the generation of antisense RNA and DNA. PRO nucleic acid sequences
CC	will also be useful for the preparation of PRO polypeptides, especially
CC	by recombinant techniques.
SQ	Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other:
Query Match	46.9%; Score 107.4; DB 21; Length 3265;
Best Local Similarity	66.8%; Pred No. 4.9e-27;
Matches 153: Conservative	0; Mismatches 76; Indels 0; Gaps 0
OY	1 GGCACAGTATCCTGGACGACACCCTGGAAAGACACTTTGTTCATCACCTTGACA 60
Db	1537 gacactgcataatgatagtagtacagcyggaagaagacagtcttcctcatcacatgyaac 1596
OY	61 AGCGACGCTGCCCAATCTCTCTCTGGATTCCAGTGGACAGAAAGCATGGCTTTGTA 120
Db	1597 agctcgtcccccagatttctctctggatcccaagtyggaacaataatgaaaattccaca 1656
OY	121 GTTGACAAAAAACCAAAATGCGCTTCCTCCAATCCACAGCATTTGCTAAGGTTGGCACT 180
Db	1657 gtgtagtgcaacttccaataatgacctatctcagatgccaggaaactgcaaagtcyggcact 1716
OY	181 TGGAAATPACGTCGTGCAAGCAGCTCACAACTTGACCTTGACTGTCA 229
Db	1717 tgggcatcacatctccaagccaagaagaccaccaaatctaactatta 1765
RESULT 15	
AAF92092	
ID	AAF92092 standard; cDNA; 3265 BP.
XX	AAF92092;
AC	15-MAY-2001 (first entry)
DT	Human PRO1124 cDNA.
XX	

```

XX Human; PRO protein; mapping; ss.
XX
XX Homo sapiens;
XX
XX WO200116318-A2.
XX
XX 08-MAR-2001.
XX
XX 24-AUG-2000; 2000WO-US23328.
XX
XX 01-SEP-1999; 99WO-US20111.
XX
XX 15-SEP-1999; 99WO-US21090.
XX
XX 07-DEC-1999; 99US-0169495.
XX
XX 05-DEC-1999; 99US-0170262.
XX
XX 11-JAN-2000; 2000US-0175481.
XX
XX 18-FEB-2000; 2000WO-US04341.
XX
XX 18-FEB-2000; 2000WO-US04342.
XX
XX 22-FEB-2000; 2000WO-US04414.
XX
XX 01-MAR-2000; 2000US-0187202.
XX
XX 03-MAR-2000; 2000US-0187202.
XX
XX 25-APR-2000; 2000US-0199397.
XX
XX 22-MAY-2000; 2000WO-US14042.
XX
XX 05-JUN-2000; 2000US-0209832.
XX
XX
XX (GENTECH ) GENENTECH INC.
XX
XX Eason DL, Filvarroff E, Gerritsen ME, Goddard A, Godowski PJ;
XX
XX Grimaldi CJ, Gunney AL, Watanabe CK, Wood RT;
XX
XX WPI; 2001-183260/18.
XX
XX P-SDB; AAB87560.
XX
XX
XX Eighty four nucleic acids encoding PRO polypeptides, useful in
XX
XX molecular biology, including use as hybridization probes, and in
XX
XX chromosome and gene mapping.
XX
XX
XX Claim 2; Fig 69; 278bp; English.
XX
XX
XX The present sequence is the coding sequence for a human PRO polypeptide
XX
XX (secreted and transmembrane). The PRO protein, and PRO agonists, PRO
XX
XX antagonists or anti-PRO antibodies are useful for preparation of a
XX
XX medicament useful in the treatment of a condition which is responsive to
XX
XX the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
XX
XX protein may also be employed as molecular weight markers for protein
XX
XX electrophoresis. The PRO coding sequence has applications in molecular
XX
XX biology, including use as hybridisation probes, and in chromosome and
XX
XX gene mapping.
XX
XX
XX Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;
XX
XX
XX
XX
XX Query: Match 46.9%; Score 107.4; DB 22; Length 3265;
XX
XX Best Local Similarity 66.8%; Pred. No. 4, 9e-27;
XX
XX Matches 153; Conservative 0; Mismatches 76; Indels 0; Gaps
XX
XX
XX 1 GGCAGGTGATCGTGACAGCACCCGTGGGAAGAGACATTGTTTCTATCACCCTGGACA 60
XX
XX 11111111111111111111111111111111111111111111111111111111
XX
XX 1537 gacactgcataatgataagtagacagtgaggaaagacacgcttcttcacacatggaac 1596
XX
XX
XX 61 ACGAGAGCTCCGCCAAATCTCTCTCTGGATCCAGTTCAGACAGAGAAGCAAGCTGGCTTTGTA 120.
XX
XX 1111111111111111111111111111111111111111111111111111111
XX
XX 1537 agctgcctccacagatattctctctggaaccacagtggaacataatggaabaattccaca 1656
XX
XX
XX 121 GTGAGCAAAAACACCAAAATGGCCCTACCTCCAAATCCAGGACATTGCTAAAGTTGGCACT 180
XX
XX 1111111111111111111111111111111111111111111111111111111
XX
XX 1657 gtgtagtcaaatccaaatgacctatctcagatattccaggaactgaaagtgggcact 1716
XX
XX
XX 181 TGGAAATACAGCTTCGCAAGCAAGCTACAAAACCTTGAACCTCGACTGCA 229
XX
XX 1111111111111111111111111111111111111111111111111111111
XX
XX 1717 tgggcattcaatcttcaagccaagcgagacaccagaaacattaaactatca 1765

```


Thu Apr 4 09:27:24 2002

Search completed: April 4, 2002, 03:17:59
Job time: 60231 sec

us-09-049-696-10.rng

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:01:12 ; Search time 22700.8 Seconds

(without alignments)
166,419 Million cell updates/sec

Title: US-09-049-696-10

Perfect score: 229
Sequence: 1 GGACAGATGATCCTGGACAG.....AACCTTGACCCCTGACTGTCA 229

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank1: *
1: gb_ba: *
2: gb_hg: *
3: gb_in: *
4: gb_lm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_lm: *
20: em_ov: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_hggo_hum: *
31: em_hggo_inv: *
32: em_hggo_rnd: *
33: em_hggo_hum: *
34: em_hggo_inv: *
35: em_hggo_rnd: *
36: em_hggo_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	229	100.0	2826	9	AK024970	AK024970 Homo sapi
2	229	100.0	2826	9	AF127036	AF127036 Homo sapi
3	229	100.0	3311	6	AX193489	AX193489 Sequence
4	229	100.0	3311	6	AF039400	AF039400 Homo sapi
5	173.6	75.8	35278	9	AF039401	AF039401 Homo sapi
6	173.6	75.8	164891	2	AL358950	AL358950 Homo sapi
7	172	75.1	113764	9	HS1651E10	AL122002 Human DNA
8	161.8	70.7	2937	10	AB017156	AB017156 Mus muscu
9	119.8	52.3	3079	4	AF095584	AF095584 Sus scrof
10	107.4	45.9	1895	9	AK000138	AK000138 Homo sapi
11	107.4	46.9	3204	9	AF127035	AF127035 Homo sapi
12	107.4	46.9	3221	9	AK000072	AK000072 Homo sapi
13	107.4	46.9	3265	6	AK092338	AK092338 Sequence
14	72.6	31.7	140718	2	AL356270	AL356270 Homo sapi
15	63.8	27.9	3288	4	AF001263	AF001263 Bos tauru
16	63.8	27.9	3288	4	AF001262	AF001262 Bos tauru
17	63.8	27.9	3317	4	AF001261	AF001261 Bos tauru
18	60.6	26.5	2984	4	BT036445	U36445 Bos tauru
19	60.6	26.5	3058	10	AF108501	AF108501 Mus muscu
20	59	25.8	2765	10	AF115852	AF115852 Mus muscu
21	59	25.8	3137	10	BC008147	BC008147 Mus muscu
22	53.6	23.4	3022	10	AF047838	AF047838 Mus muscu
23	53.6	23.4	3471	10	AF052746	AF052746 Mus muscu
24	52.6	23.0	3415	9	AF043976	AF043976 Homo sapi
25	43.2	18.9	140718	2	AL356270	AL356270 Homo sapi
26	43.2	18.9	175591	2	AC068071	AC068071 Homo sapi
27	37.2	16.2	147292	2	AC023010	AC023010 Homo sapi
28	37.2	16.2	211393	2	AL513347	AL513347 Mus muscu
29	37.2	16.2	216239	2	AL513356	AL513356 Mus muscu
30	35	15.3	124775	2	AC010653	AC010653 Homo sapi
31	35	15.3	125768	2	AC026451	AC026451 Homo sapi
32	35	15.3	199450	9	AC005674	AC005674 Homo sapi
33	34.2	14.9	2832	9	AF127980	AF127980 Homo sapi
34	34.2	14.9	2970	9	AF043977	AF043977 Homo sapi
35	34.2	14.9	3604	9	AB026833	AB026833 Homo sapi
36	34.2	14.9	4077	6	AX054697	AX054697 Sequence
37	34.2	14.9	124096	2	AL354988	AL354988 Homo sapi
38	34.2	14.9	136857	2	AC091600	AC091600 Homo sapi
39	33.6	14.7	33770	3	CEB154	CEB154
40	33.4	14.6	106742	9	AC008904	AC008904 Homo sapi
41	33.4	14.6	128702	9	AC008872	AC008872 Homo sapi
42	33.2	14.5	10655	14	AF289029	AF289029 denque vi
43	33	14.4	46556	2	AC069464	AC069464 Homo sapi
44	33	14.4	151421	9	AL391865	AL391865 Human DNA
45	33	14.4	199042	2	AC068299	AC068299 Homo sapi

ALIGNMENTS

RESUME 1
AK024970
LOCUS Homo sapiens cDNA: FLJ21317 fls, clone COL02275, highly similar to
DEFINITION AF127036 Homo sapiens calcium-activated chloride channel protein 1
(CACCL1) mRNA.
ACCESSION AK024970
VERSION AK024970.1 GI:10437397
KEYWORDS Oligo cloning; fls (full insert sequence);
SOURCE Homo sapiens colon cDNA to mRNA, clone lib:COL clone:COL02275.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 2022)

AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, T.

TITLE Direct Submission

JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: 5'-6' end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source 1..2022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL02275"
/clone_11b="COL"
/issue_type="colon"
/note="cloning vector pME18SFL3"
1..2022
/note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA"

misc_feature 612 a 472 c 453 g 485 t

BASE COUNT 612 a 472 c 453 g 485 t

ORIGIN

Query Match 100.0%; Score 229; DB 9; Length 2022;
Best Local Similarity 100.0%; Pred. No. 2,3e-68;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCACAGTGCATCTGGACAGCAGCGTGGAAAGACACTTTGTTCTATCACCCTGGACA 60
|||||
DB 723 GGCACAGTGCATCTGGACAGCAGCGTGGAAAGACACTTTGTTCTATCACCCTGGACA 782
OY 61 ACGCAGCCTCCCAAAATCTCTCTGGGATCCAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 120
|||||
DB 783 ACGCAGCCTCCCAAAATCTCTCTGGGATCCAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 842
OY 121 GTGGACAAAACACCAAAATGCGCTACCTCCAAATCCAGGCAATGCTTAAGTTGGCACT 180
|||||
DB 843 GTGGACAAAACACCAAAATGCGCTACCTCCAAATCCAGGCAATGCTTAAGTTGGCACT 902
OY 181 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGAACCTTGAGACTGCA 229
|||||
DB 903 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGAACCTTGAGACTGCA 951

RESULT 2

AF127036 2826 bp mRNA PRI 10-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)
DEFINITION mRNA, complete cds.
ACCESSION AF127036
VERSION AF127036.1 GI:4585468
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 2826)
AUTHORS Agnel, M., Vernat, T. and Culouscou, J.-M.
TITLE Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea
JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792
REFERENCE 2 (bases 1 to 2826)

AUTHORS Agnel, M. and Culouscou, J.-M.

TITLE Direct Submission

JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrieres, Ruell-Malmaison 92500, France

FEATURES

source 1..2826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="small intestine; colon"
1..2826
/gene="CaCC1"
1..2826
/gene="CaCC1"
5..2749
/gene="CaCC1"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 1"
/protein_id="A025487.1"
/db_xref="GI:4585469"

BASE COUNT 875-a 623 c 632 g 696 t

ORIGIN

Query Match 100.0%; Score 229; DB 9; Length 2826;
Best Local Similarity 100.0%; Pred. No. 2,5e-68;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCACAGTGCATCTGGACAGCAGCGTGGAAAGACACTTTGTTCTATCACCCTGGACA 60
|||||
DB 1544 GGCACAGTGCATCTGGACAGCAGCGTGGAAAGACACTTTGTTCTATCACCCTGGACA 1573
OY 61 ACGCAGCCTCCCAAAATCTCTCTGGGATCCAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 120
|||||
DB 1574 ACGCAGCCTCCCAAAATCTCTCTGGGATCCAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 1633
OY 121 GTGGACAAAACACCAAAATGCGCTACCTCCAAATCCAGGCAATGCTTAAGTTGGCACT 180
|||||
DB 1634 GTGGACAAAACACCAAAATGCGCTACCTCCAAATCCAGGCAATGCTTAAGTTGGCACT 1693
OY 131 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGAACCTTGAGACTGCA 229
|||||
DB 1694 TGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGAACCTTGAGACTGCA 1742

RESULT 3

AX193489 3311 bp DNA PAT 15-AUG-2001
LOCUS Sequence 1056 from Patent WO0149716.
DEFINITION AX193489
ACCESSION AX193489
VERSION AX193489.1 GI:15211440
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 3311)
AUTHORS Xu, J., Lodes, M.J., Secorist, H., Benson, D.R., Meagher, M.J.,

TITLE Stolk, J.A., King, G.E., Wang, T. and Jiang, Y.
Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
Patent: WO 0149716-A 1056 12-JUL-2001;
JOURNAL CORIAX CORPORATION (US)

FEATURES
source location/Qualifiers
1..3311
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 100.0%; Score 229; DB 6; Length 3311;
Best local Similarity 100.0%; Pred. No. 2.6e-68;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACAGTGTGCTGGACAGCACCCTGGGAAGACACTTGTTCATCACCCTGGACA 60
DB 1861 GGCACAGTGTGCTGGACAGCACCCTGGGAAGACACTTGTTCATCACCCTGGACA 1920
QY 61 ACCGAGCCTCCCAATCCTCTCTGGATCCAGTGGACAGCAAGGAGGCTTGTGA 120
DB 1921 ACCGAGCCTCCCAATCCTCTCTGGATCCAGTGGACAGCAAGGAGGCTTGTGA 1980
QY 121 GTGGACAAAAACACCAAAATGGCTTACCTCAATCCAGGCAATGCTTAAGTTGGCACT 180
DB 1981 GTGGACAAAAACACCAAAATGGCTTACCTCAATCCAGGCAATGCTTAAGTTGGCACT 2040
QY 181 TGGAAATACAGTCTGCAGCAAGCTCAAAACCTTGACCTGACTGTCA 229
DB 2041 TGGAAATACAGTCTGCAGCAAGCTCAAAACCTTGACCTGACTGTCA 2089

RESULT 4
AF039400 3311 bp mRNA PRI 14-DEC-1998
LOCUS Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA,
DEFINITION complete cds.
ACCESSION AF039400
VERSION AF039400.1 GI:4009457
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3311)
AUTHORS Gruber, A.D., Elble, R.C., Ji, H.L., Schreur, K.D., Fuller, C.M. and Paul, B.U.
TITLE Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 3311)
AUTHORS Gruber, A.D., Elble, R. and Paul, B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

FEATURES
source location/Qualifiers
1..3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/tissue.type="small intestine"
1..3311
/gene="hCLCA1"
352..3096
/gene="hCLCA1"
/note="transmembrane glycoprotein"
/codon.start=1
/product="calcium-dependent chloride channel-1"

gene
CDs

/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPFKSSVFLILHLLEGALNSLIQNNNGEYIVADPNP
EDTLIOIKDVTQASIVLEFATGRKRVNVAIIPETWKRAVYRKLTYNA
DIVVAESTPRGNDEPTQONGNGEGERHLLPDLIAGKKLAEYPOGAFYENAH
LRMGVDEINDEKFTLSNGRIVAVRCSAGITGVNKKCOGSCYTRKRVGL
YKGECEVLOSROTERKASIMFAOHVDSIVFCEPHNNKRAPKQOKNRSTWEI
RDESEFKTTPMTOPNPFFSLIOGORVLCVILDKSGSMATGNRLNRLNOAGLFL
LOTVEGSGWGTVPDPAHVOSELIOINGSDRDLARKLPAAAGGTSICGSLSA
FYIIRKKYPPDGSSEIVLLDGEENITSGCNEVKOGAIIHYVALPSAOLEEELSK
MGGLIOTYASDOYQNNGLIDAFALSSGNKAVRSIQLESKLTJONSOMNGITYIV
DSTVGKDTLFLITWTPQPIILMDPSGOQGFVVDKNTKMAVLIQIGIAKVTWKY
SLQASSQTLTLVTLSRASNATLPPIVTSKNTDPSKPPSLVYVANIIRGASPIIRA
SVYALIESVNGKVTLELDNGAGADATKDGVSRYFTTYDNGRYSVRAIGLVN
AARRVYIPQOOSGALYIPGMEINDEIOMNPREIINDDOYHOKVCRSPSSGGSFVAS
DVNAPRIPLDFPFGQITDLKATFHGSLITLWTAGDDVDHTTAFKRYIIRISTSID
LRKFNESLQVNTTALIPKANESEVFLPRENITENGTDLFIATQAVDKVLSKSEI
SNIAVSLFIPQTPPEPSPDETSAPCPNIRHNSIPGIIHLIKMKWIGELQLSIA

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 100.0%; Score 229; DB 9; Length 3311;
Best local Similarity 100.0%; Pred. No. 2.6e-68;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACAGTGTGCTGGACAGCACCCTGGGAAGACACTTGTTCATCACCCTGGACA 60
DB 1861 GGCACAGTGTGCTGGACAGCACCCTGGGAAGACACTTGTTCATCACCCTGGACA 1920
QY 61 ACCGAGCCTCCCAATCCTCTCTGGATCCAGTGGACAGCAAGGAGGCTTGTGA 120
DB 1921 ACCGAGCCTCCCAATCCTCTCTGGATCCAGTGGACAGCAAGGAGGCTTGTGA 1980
QY 121 GTGGACAAAAACACCAAAATGGCTTACCTCAATCCAGGCAATGCTTAAGTTGGCACT 180
DB 1981 GTGGACAAAAACACCAAAATGGCTTACCTCAATCCAGGCAATGCTTAAGTTGGCACT 2040
QY 181 TGGAAATACAGTCTGCAGCAAGCTCAAAACCTTGACCTGACTGTCA 229
DB 2041 TGGAAATACAGTCTGCAGCAAGCTCAAAACCTTGACCTGACTGTCA 2089

RESULT 5
AF039401 35278 bp DNA PRI 14-DEC-1998
LOCUS Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene,
DEFINITION complete cds.
ACCESSION AF039401
VERSION AF039401.1 GI:4009459
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 35278)
AUTHORS Gruber, A.D., Elble, R.C., Ji, H.L., Schreur, K.D., Fuller, C.M. and Paul, B.U.
TITLE Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 35278)
AUTHORS Gruber, A.D., Elble, R. and Paul, B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Department of Pathology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

FEATURES
source location/Qualifiers
1..35278
/organism="Homo sapiens"
/db_xref="taxon:9606"

gene
CDs


```

/promoter
  /chromosome="1"
  /map="1p22-p31"
  1.1617
  /gene="hCLCA1"
  1.33522
  /gene="hCLCA1"
  767.1265
  /rpl_family="r1 transposable element"
  /rpl_type="dispersed"
  1587.1596
  /gene="hCLCA1"
  /note="putative"
  1618.1708
  /gene="hCLCA1"
  /number=1
  join(1618..1708,1962..2383,6667..6807,6979..7126,
  9696..9801,15454..15631,18591..18809,19783..20010,
  22253..22427,24522..24628,26644..26859,27447..27708,
  28766..28936,31840..32079,32919..33522)
  /gene="hCLCA1"
  /product="calcium-dependent chloride channel-1"
  1962..2383
  /gene="hCLCA1"
  /number=2
  join(2222..2383,6667..6807,6979..7126,9696..9801,
  15454..15631,18591..18809,19783..20010,22253..22427,
  24522..24628,26644..26859,27447..27708,28766..28936,
  31840..32079,32919..33510)
  /gene="hCLCA1"
  /codon_start=1
  /product="calcium-dependent chloride channel-1"
  /protein_id="AAC95429.1"
  /gb_xref="GI:4009460"
  /translation="MGPKRSYFILLIHLEAGLSNLIOLNNNGEIVAVIDPNV
  EDELIIQIKDWTQASLLEPATKRRFFKVALILIPETKTKADYRPRKIETKNA
  DLVAVSEPPGDEPYEQMGNGGKGEIHLPFLPGLAEFGPKAIVHMAH
  LRMVGEDEYNDEKFLYSNGRIQAVRCSGIGTVNKKCGSGCYTRKCFNKVGL
  YEKGEFVLOSROTERKASIMFAQHVDISYEFTEDONHNEAPNOKNCLNTEWVI
  RSEDEKFTPTPTOPNPPTFSLQIGRIYCLVDKSGSMATGNRLNINLQGLFL
  IOTVELGSWVGWTFEDSAHQVSELIOINSGSDRDLAKRLPAASGGSISGKSLA
  PTVIKKYPTDSEIYLLTDGDNLTISGDNVKGSGAIITHTVALGPSAOGLEELSK
  MTGGLDTASDOVONGLIDAFGALSNGGAVSQSIOLESGLTLONSOMNGEIVY
  DSTVSKDILFTLTWTPQIILLMPDSGKOGFVADKTKKAVIQLPEIAVGVWKY
  SLQASQTLITLITSNASNATLPITVYSKTRKIDSKPEPPLVAVANIRQGSPLIRA
  SYTALIESVNGKTVLELDNGAGADATDDCVSRKFTYDNGRSYKVALGAGV
  AARRVYPOGSCALYIPGWIENDEIOMNPDEINKDDVOHQVCFSRSSGSPVAS
  DYPNAPIPDLPPGOTIDNARIHGGSILINLTWAPGDODYHAKYIRISTSLID
  LDKRNEISQVWTALITPEANSEVELEKPEPTENCTDIFILQAVDRKDIKSEI
  SNIRVSLFIPQTPPTPSPDTSAPCFNHNINSTITPQIHILKIMKWTGELQSLIA
  "
  exon
    6667..6807
    /gene="hCLCA1"
    /number=3
    6979..7126
    /gene="hCLCA1"
    /number=4
    9696..9801
    /gene="hCLCA1"
    /number=5
    15454..15631
    /gene="hCLCA1"
    /number=6
    18591..18809
    /gene="hCLCA1"
    /number=7
    19783..20010
    /gene="hCLCA1"
    /number=8
    22253..22427
    /gene="hCLCA1"
    /number=9
    24522..24628
    /gene="hCLCA1"
  exon

```

```

exon
  /number=10
  26644..26859
  /gene="hCLCA1"
  /number=11
  27447..27708
  /gene="hCLCA1"
  /number=12
  28766..28936
  /gene="hCLCA1"
  /number=13
  31840..32079
  /gene="hCLCA1"
  /number=14
  32919..33522
  /gene="hCLCA1"
  /number=15
  exon
  BASE COUNT 11245 a 6870 c 7230 g 9933 t
  ORIGIN

```

```

Query Match 75.8% Score 173.6; DB 9; Length 35278;
Best Local Similarity 97.8%; Pred. No. 7.4e-49;
Matches 176; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 GGCACAGTATGATGTGACAGCAGCGTGGAAAGACACTTGTCTTATCAGCTGACAC 60
  |||||
Db 26689 GGCACAGTATGATGTGACAGCAGCGTGGAAAGACACTTGTCTTATCAGCTGACAC 26748
  |||||
QY 61 AGCGACCTCCCAATCTCTTCTGCGATCCAGTGGACACACCAAGCTGCTTGTGA 120
  |||||
Db 26749 AGCGACCTCCCAATCTCTTCTGCGATCCAGTGGACACACCAAGCTGCTTGTGA 26808
  |||||
QY 121 GTGACAAACACCAAAATGGCTACCTCCAAATCCAGCATTTGTAAGTGGCACT 180
  |||||
Db 26859 GTGACAAACACCAAAATGGCTACCTCCAAATCCAGCATTTGTAAGTGGCACT 26868
  |||||

```

```

RESULT 6
AL358950/c
LOCUS AL358950
DEFINITION Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN
PROGRESS ***. 10 unordered pieces.
ACCESSION AL358950
VERSION AL358950.4 GI:12539689
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
  1 (bases 1 to 164891)
  Plumb,B.
  Direct Submission
  Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
  requests: clonequest@sanger.ac.uk
  On Jan 26, 2001 this sequence version replaced gi:9988471.
  COMMENT
    ----- Genome Center
    Center: Sanger Centre
    Center code: SC
    Web site: http://www.sanger.ac.uk
    Contact: humquery@sanger.ac.uk
    Project Information
    Center project name: BA444C12
    ----- Summary Statistics
    Assembly program: XGAP4; version 4.5
    Sequencing vector: plasmid; 108752; 100% of reads
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Consensus quality: 161347 bases at least Q40
    Consensus quality: 162610 bases at least Q30
    Consensus quality: 163328 bases at least Q20
    Insert size: 163991; sum-of-ctrls
    Insert size: 163373; 6.6% error; agarose-fp
    Quality coverage: 5.19x in Q20 bases; sum-of-ctrls Quality

```


Coverage: 5.21x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 34236: contig of 34236 bp in length
34237 34336: gap of 100 bp
34337 47372: contig of 13036 bp in length
47373 47472: gap of 100 bp
47473 67025: contig of 19553 bp in length
67026 67125: gap of 100 bp
67126 71239: contig of 4114 bp in length
71240 71339: gap of 100 bp
71340 83903: contig of 12564 bp in length
83904 84003: gap of 100 bp
84004 91357: contig of 7354 bp in length
91358 91457: gap of 100 bp
91458 96452: contig of 4995 bp in length
96453 96552: gap of 100 bp
96553 148189: contig of 51637 bp in length
148190 148289: gap of 100 bp
148290 159778: contig of 11489 bp in length
159779 159878: gap of 100 bp
159879 164891: contig of 5013 bp in length.

```

FEATURES

source

```

1.164891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone_1lb="RPCT-11.2"
1.34236
/note="assembly-fragment:00027
fragment_chain:1
clone_end:77
vector_side:left"
misc_feature
34337..47372
/note="assembly-fragment:00312
fragment_chain:1"
47473..67025
/note="assembly-fragment:01923
fragment_chain:1"
67126..71239
/note="assembly-fragment:02094
fragment_chain:1"
71340..83903
/note="assembly-fragment:02213
fragment_chain:1"
84004..91357
/note="assembly-fragment:01292
fragment_chain:1"
91458..96452
/note="assembly-fragment:01391
fragment_chain:1"
96553..148189
/note="assembly-fragment:02045
fragment_chain:1"
148290..159778
/note="assembly-fragment:00427
fragment_chain:1"
159879..164891
/note="assembly-fragment:00773
fragment_chain:1"
misc_feature
159879..164891
/note="assembly-fragment:00773
fragment_chain:1"
clone_end:SP6
vector_side:right"
BASE COUNT 48381 a 32179 c 32357 g 51067 t 907 others
ORIGIN

```

Query=Match 75.8%; Score 173.6; DB 2; Length 164891;

Beg=Local Similarity 97.8%; Pred. No. 1e-48; Matches 176; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

DB 55115 GGCACATGATCGTGACACACCGTGGAAGACACTTGTCTTATCACCCTGGACA 55256

Query Match 75.1%; Score 172; DB 9; Length 113764;
 Best Local Similarity 97.2%; Pred. No. 3,4e-48;
 Matches 175; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DB 26399 GGCACAGTATCGTGACAGCACCCTGGGAAAGACACTTGTCTTATACCTGGACA 60
 1 GGCACAGTATCGTGACAGCACCCTGGGAAAGACACTTGTCTTATACCTGGACA 60
 61 ACCGACCTCCCAATCTCTCTGGATCCAGTGAGACAGAGGAGGCTTTGTA 120
 DB 26459 ATGCACACCTCCCAATCTCTCTGGATCCAGTGAGACAGAGGAGGCTTTGTA 26518
 121 GTGCACAAACACCAAAATGGCTACCTCAATCCAGGCTTGTCTAAGGTTGGCACT 180
 DB 26519 GTGCACAAACACCAAAATGGCTACCTCAATCCAGGCTTGTCTAAGGTTGGCACT 26578

RESULT 8

AB017156 2937 bp mRNA ROD 10-NOV-1999
 LOCUS Mus musculus gob-5 mRNA, complete cds.
 DEFINITION AB017156
 ACCESSION AB017156.1 GI:3721911
 VERSION
 KEYWORDS
 SOURCE Mus musculus adult intestine goblet cell cDNA to mRNA.
 ORGANISM

REFERENCE 1 (sites)
 AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.
 TITLE Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice
 JOURNAL Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)
 MEDLINE 99160866
 REFERENCE 2 (bases 1 to 2937)
 AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.
 TITLE Direct Submission
 JOURNAL Submitted (26-AUG-1998) to the DDBJ/EMBL/GenBank databases, Tohru Komiyama, ERATO, JST, Genome Asymmetry Group, Doi Biosymmetry Project, Tokodai 5-9-9, Tsukuba, Ibaragi 300-2635, Japan
 (E-mail: tkom@bioa.eriato.trc-net.co.jp, Tel: 81-298-48-1515, Fax: 81-298-47-8901)

FEATURES

1..2937 location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_type="goblet cell"
 /dev_stage="adult"
 /tissue_type="intestine"
 15..2756 /gene="gob-5"
 15..2756 /gene="gob-5"
 15..2756 /gene="gob-5"
 /codon_start=1
 /protein_id="BAA3743.1"
 /db_xref="GI:3721912"
 /translation="MESLKPSPVLLILHLEGVLSLSLQJLNNGYEGVIAIDHDP
 EDRLIOHIDMTQASPYLFEATGKRFYRKNVAIIPESMRAPETRTKLETFKNA
 DYLVSTSPGNDPEYREHAGGEGKIRHLHPDLGAKKLYOYGODTFEHEHVAH
 FRWCVEYENDEKFEYSKGRPOAVRCSAITGKNOVRCGGSCITNGCVIDRATG
 LYKNCVFPDPHONERKASIMFNONINSYVEFCEKHNHONBANDONONONASTMEV
 IOESDEKQTPMTAPAPAPFSLQIGRIICVLDKSGSMLNDRNLNMANASRKF
 LAOTVEGSGWGWGTPDSAPYVSELKOLNSGADRLLHLPTVSAAGTSCISGLRT
 AFYIKKKYPTDGESEILVDGEDNTISSCFDAIKOSGAIHIVALPAPAKLELDIS
 KMGGLQTSDDVONNGVDAFALSSGNAALAOHSIOLESRGVNLONNMGNSYI
 VDSVSGDLEFLITWTHPTPIITWDSGVENGEFLDITTKAVYLOVPGTAYGEWK
 YSIQASSOTLTLVTSRAASATLPTITVTYVYVKNCKGPPSPVTVVASISOGASPLR
 ASVATLIESVNGKVTLELDNGAGADATKWDGYSRFTAFDANGYRSTKYNAGGV
 TSDORAPAPKRNAMYIDGWTEDVEFRNRPRTSYVDQDCEFSRTSSGSEVARN
 VPAAPIDLPFPCCOITDLKASIOGNLVLNLTWAPGDDYDHGASNVILRMSTSID
 LRHFNNTSLGNTGLIPKEASSEEIEFEELGNTFGNTDIFLAIOAVDKSNLKSSEI

gene

CDS

BASE COUNT

860 a 718 c 693 g 666 t
 SNIRSVFIPAPDEPPEPDPDSTPPCPDISINSTIPGIIHLKIMKMLGEMQVTLGLH"

ORIGIN

Query Match 70.7%; Score 161.8; DB 10; Length 2937;
 Best Local Similarity 81.7%; Pred. No. 5,4e-45;
 Matches 187; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

DB 1537 GGCACAGTATCGTGACAGCACCCTGGGAAAGACACTTGTCTTATACCTGGACA 60
 1 GGCACAGTATCGTGACAGCACCCTGGGAAAGACACTTGTCTTATACCTGGACA 60
 61 ACCGACCTCCCAATCTCTCTGGATCCAGTGAGACAGAGGAGGCTTTGTA 120
 DB 1537 GGCACAGTATCGTGACAGCACCCTGGGAAAGACACTTGTCTTATACCTGGACA 1586
 121 GTGCACAAACACCAAAATGGCTACCTCAATCCAGGCTTGTCTAAGGTTGGCACT 180
 DB 1647 CTAGCACACACCACTAAGGTTGGCTACCTCCAGGCTTGTCTAAGGTTGGCTTT 1706
 131 TCGAAATGACCTTCGCAAGCAGCTCAACCTTGACCTGACTGTCA 229
 DB 1767 TCGAAATGACCTTCGCAAGCAGCTCAACCTTGACCTGACTGTCA 1755

RESULT 9

AF095584 3079 bp mRNA MM 13-OCT-2000
 LOCUS Sus scrofa epithelial chloride channel protein (AEC) mRNA,
 DEFINITION AF095584
 ACCESSION AF095584.1 GI:6002645
 VERSION
 KEYWORDS
 SOURCE pig, cv
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE 1 (bases 1 to 3079)
 AUTHORS Gaspar, K.J., Racette, K.J., Gordon, J.R., Loewen, M.E. and Forsyth, G.W.
 TITLE Cloning a chloride conductance mediator from the apical membrane of porcine ileal enterocytes
 JOURNAL Physiol. Genomics (Online) 3 (2), 101-111 (2000)
 MEDLINE 20473747
 REFERENCE 2 (bases 1 to 3079)
 AUTHORS Gaspar, K.J., Gabriel, S.E., Racette, K.J. and Forsyth, G.W.
 TITLE Direct Submission
 JOURNAL Submitted (28-SEP-1998) Veterinary Physiological Sciences, University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4, Canada

FEATURES

1..3079 location/Qualifiers
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /tissue_type="ileal mucosa"
 1..3079 /gene="AEC"
 133..2886 /gene="AEC"
 /function="Involved in exocrine chloride conductance"
 /codon_start=1
 /product="epithelial chloride channel protein"
 /protein_id="AAF00077.1"
 /db_xref="GI:6002645"
 /translation="MGRSSRFITVILHLEGAOSNSLIQNGNGEYVIAIDPNV
 EDRLIOHIDMTQASPYLFEATGKRFYRKNVAIIPESMRAPETRTKLETFKNA
 DYLVSTSPGNDPEYREHAGGEGKIRHLHPDLGAKKLYOYGODTFEHEHVAH
 FRWCVEYENDEKFEYSKGRPOAVRCSAITGKNOVRCGGSCITNGCVIDRATG
 LYKNCVFPDPHONERKASIMFNONINSYVEFCEKHNHONBANDONONONASTMEV
 IOESDEKQTPMTAPAPAPFSLQIGRIICVLDKSGSMLNDRNLNMANASRKF
 LAOTVEGSGWGWGTPDSAPYVSELKOLNSGADRLLHLPTVSAAGTSCISGLRT
 AFYIKKKYPTDGESEILVDGEDNTISSCFDAIKOSGAIHIVALPAPAKLELDIS
 KMGGLQTSDDVONNGVDAFALSSGNAALAOHSIOLESRGVNLONNMGNSYI
 VDSVSGDLEFLITWTHPTPIITWDSGVENGEFLDITTKAVYLOVPGTAYGEWK
 YSIQASSOTLTLVTSRAASATLPTITVTYVYVKNCKGPPSPVTVVASISOGASPLR
 ASVATLIESVNGKVTLELDNGAGADATKWDGYSRFTAFDANGYRSTKYNAGGV
 TSDORAPAPKRNAMYIDGWTEDVEFRNRPRTSYVDQDCEFSRTSSGSEVARN
 VPAAPIDLPFPCCOITDLKASIOGNLVLNLTWAPGDDYDHGASNVILRMSTSID
 LRHFNNTSLGNTGLIPKEASSEEIEFEELGNTFGNTDIFLAIOAVDKSNLKSSEI

gene

CDS

CDS

/gene="CaCC2"
29..2782
/gene="CaCC2"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 2"
/protein_id="AA048398.1"
/db_xref="gi:5726289"
/translation="MGLEPGFVLLVCLLHOSNTSIFIKLNNGFEDIVIVDPSE
DEKILIEQEDVATSYLIPFATKREFFKNSVLLIENKNEQYRPRHEHKAH
VIVAPFLTGEDVPTKQPTKCEKGRVILHTDILGKONFEGPKLFEHMAH
RMGVFDEYNEDQPPYRAKSKTEATRSAGISGNRRYKQGGCSLRARIDSTTL
YKDCQEPEDVPTKQPTKCEKGRVILHTDILGKONFEGPKLFEHMAH
SNSDEKTIIPATVPPEPVSILKIRQIVCLVLDKSGSGMGRDLNRMQAKHFL
LOVNGSWGVAMHEDSPATLVNKLIOIKSDEBENTIMAGLPTPLGTSICGDKYA
FOVIGELSHOLDSEVILLTDGEDNTASCIIDEKOSGAIYHIALGRADAVTEMS
KITGSHHYVDEDMONNGLIDAFGLSGNTDLGKSLQLEKGLTINSNAMNDYVI
IDSTVKGDFEFLITNSLPSISLMDPSGTTMEMPYDASKMAVLSIPGAKYGVAV
YNLQAKNPEFLITNSLPSISLMDPSGTTMEMPYDASKMAVLSIPGAKYGVAV
LGANTATIESONGHEVLELLDNGAGDSRKNQGVSRITATENGSRYSLKRAH
GANTARLKRLPPLNRAAYIPGVVNGELEAMPREPIDEDQTELEDSRTAGAV
VSQVSLPLPDQPPSQTLDATVHEKILITWAPEDNDFGVKORYILIRISALIL
DIRSDPDALQVNTDLSPEKNSKESAPFENISENATHIFILKSIDKSNITSK
VSNIAQVILFIPQAPDIDIPPTPPIDKSHNGVISTVLVIGSVIVAEILLS
TTL"

BASE COUNT 1098 a 594 c 633 g 879 t
ORIGIN

Query Match 46.9%; Score 107.4; DB 9; Length 3204;
Best Local Similarity 66.8%; Pred. No. 4.2e-26;
Matches 153; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1 GGCACAGTATCGTGACAGACCGTGGAAGACACTTTGTTTATCACCCTGGACA 60
DB 1541 GACACGTGCAATATGATAGTACAGTGGGAAGACAGCTTCTTTCACACATGGAAC 1600
QY 61 ACGAGAGCTCCCAAAATCTCTCTGGATCCGATCGACAGAGAAGAGTGCTTGTGA 120
DB 1601 AGCTGCTCCCAATATCTCTCTGGATCCGATCGACAGAGAAGAGTGCTTGTGA 1660
QY 121 GTGACAAAAAACCAAAATGGCTTACCTCCAAATCCAGGACATGCTAAGTTGGCACT 180
DB 1661 GTGATGCAATTCCTCAAAATGGCTTACCTCCAAATCCAGGACATGCTAAGTTGGCACT 1720
QY 181 TGGAAATACAGTCTGACAGCAAGCTCACAAACCTTGACCCGACTGCA 229
DB 1721 TGGGATACATCTTCAGCCAAAGCAAGCCGAAACATTAATTA 1769

RESULT 12
AK000072 3221 bp mRNA PRI 22-FEB-2000
LOCUS Homo sapiens cDNA FLJ20065 fis, clone COL01613, highly similar to
DEFINITION ECUC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.
ACCESSION AK000072
VERSION AK000072.1 GI:7019922
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_11b.COL clone:COL01613.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Ohtani,T., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE
JOURNAL NEDO human cDNA sequencing project
REFERENCE
2 (bases 1 to 3221)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
AUTHORS
TITLE
JOURNAL Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio

Sugano, Institute of Medical Science, University of Tokyo, Department
of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@iims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by science and Technology
Agency).

FEATURES
source
1..3221
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01613"
/clone_lib="COL"
/issue_type="colon"
/note="cloning vector pME18SFL3"
1..3221
/note="highly similar to ECUC_BOVIN EPITHELIAL CHLORIDE
CHANNEL PROTEIN"

BASE COUNT 1105 a 600 c 634 g 882 t
ORIGIN

Query Match 46.9%; Score 107.4; DB 9; Length 3221;
Best Local Similarity 66.8%; Pred. No. 4.2e-26;
Matches 153; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1 GGCACAGTATCGTGACAGACCGTGGAAGACACTTTGTTTATCACCCTGGACA 60
DB 1555 GACACGTGCAATATGATAGTACAGTGGGAAGACAGCTTCTTTCACACATGGAAC 1614
QY 61 ACGAGAGCTCCCAAAATCTCTCTGGATCCGATCGACAGAGAAGAGTGCTTGTGA 120
DB 1615 AGCTGCTCCCAATATCTCTCTGGATCCGATCGACAGAGAAGAGTGCTTGTGA 1674
QY 121 GTGACAAAAAACCAAAATGGCTTACCTCCAAATCCAGGACATGCTAAGTTGGCACT 180
DB 1675 GTGATGCAATTCCTCAAAATGGCTTACCTCCAAATCCAGGACATGCTAAGTTGGCACT 1734
QY 181 TGGAAATACAGTCTGACAGCAAGCTCACAAACCTTGACCCGACTGCA 229
DB 1735 TGGGATACATCTTCAGCCAAAGCAAGCCGAAACATTAATTA 1783

RESULT 13
AX092338 3265 bp DNA PAT 21-MAR-2001
LOCUS Sequence 69 from Patent WO0116318.
DEFINITION AX092338
ACCESSION AX092338
VERSION AX092338.1 GI:13444483
KEYWORDS human
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3265)
Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
Wood,M.I.
TITLE
JOURNAL Secreted and transmembrane polypeptides and nucleic acids encoding
the same
Patent: WO 0116318-A 69 08-MAR-2001;
Genentech, Inc. (US)
FEATURES
source
1..3265
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1159 a 596 c 632 g 878 t
ORIGIN

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:33 ; Search time 16681 Seconds

(Without alignments)
123.041 Million cell updates/sec

Title: US-09-049-696-11

Perfect score: 191
Sequence: 1 GGCTTTGTAGTGACACAAAAA.....AACAGGACACCAACCAATT 191

Scoring table: IDENTITY NUC
Gapop 10.0 ; Gapext 1.0

Searched: 11351937 segs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gp_est1:*
11: gp_est2:*
12: gp_hic:*
13: gp_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rnd:*
20: em_gss_vrt:*
21: em_gss_other:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	191	100.0	356	10	AM361528 QV2-CT026
C 2	191	100.0	496	10	AM753448 QV2-CT026
C 3	191	100.0	633	10	AM753451 QV2-CT026
C 4	191	100.0	653	10	AM361534 QV2-CT026
5	153.4	80.3	625	10	AM361532 QV2-CT026
6	137.2	71.8	621	10	AM361520 QV2-CT026
7	131.4	68.8	657	10	AM361521 QV2-CT026
8	123.8	64.8	455	10	AA871197 QV2-CT026
9	123.8	64.8	469	10	BE136935 QV2-CT026
10	123.8	64.8	974	11	BC962464 QV2-CT026
11	123.8	64.8	2915	12	AK007466 QV2-CT026
12	123.8	64.8	2933	12	AK008659 QV2-CT026

13	120.8	63.2	652	10	AM753449 QV2-CT026
14	119.8	59.7	561	10	AM361523 QV2-CT026
15	113.4	52.4	906	11	BF581342 QV2-CT026
16	113.4	59.4	911	11	BF578693 QV2-CT026
17	113.4	59.4	935	11	BF578693 QV2-CT026
18	105.6	55.3	479	10	AA541829 QV2-CT026
19	96.4	50.5	1107	11	BF581120 QV2-CT026
20	85.8	44.9	703	13	AZ022011 RPCI-23-3
21	84.2	44.1	539	13	AZ464851 QV2-CT026
22	84	44.0	534	11	BF355513 QV2-CT026
23	75	39.3	266	10	BB072609 QV2-CT026
24	73.2	38.3	550	10	BF355533 QV2-CT026
25	65	34.0	450	10	AA734088 QV2-CT026
26	61.2	32.0	654	10	AM361522 QV2-CT026
27	60.8	31.8	308	10	AV077718 QV2-CT026
28	45.2	23.7	538	10	BE137016 QV2-CT026
29	44	23.0	774	11	BF583551 QV2-CT026
30	38.2	20.0	747	11	BF580957 QV2-CT026
31	35.8	18.7	375	10	AM324787 QV2-CT026
32	35.8	18.7	471	11	BF199173 QV2-CT026
33	34	17.8	443	13	TA33E070 QV2-CT026
34	33.4	17.5	603	11	BI233460 QV2-CT026
35	33.4	17.5	638	11	BI180267 QV2-CT026
36	33.2	17.4	776	13	AQ251017 QV2-CT026
37	32.8	17.2	807	13	CNS0128R QV2-CT026
38	32.6	17.1	676	11	BG842416 QV2-CT026
39	32.4	17.0	425	11	BF614276 QV2-CT026
40	32.4	17.0	502	10	AA757956 QV2-CT026
41	32.4	17.0	762	11	BF613755 QV2-CT026
42	32.2	16.9	386	11	T78952 QV2-CT026
43	32.2	16.9	751	11	BF677473 QV2-CT026
44	31.8	16.6	327	13	AO182324 QV2-CT026
45	31.8	16.6	704	11	BF916205 QV2-CT026

ALIGNMENTS

RESULT 1
LOCUS: AM361528/c
DEFINITION: QV2-CT0261-261099-011-c04 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION: AM361528
VERSION: AM361528.1 GI:6866282
KEYWORDS: EST.
SOURCE: human.
ORGANISM: Homo sapiens
REFERENCE: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 356)
HGCP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL.
(http://www.ludwig.org.br/scripts/gethtml2.pl?tbl=QV2&ct=QV2-CT0261-261099-011-c04&ts=1999-10-26&ta=1)
Seq primer: puc 18 forward
High quality sequence start: 356.
Location/Qualifiers
1. 356
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"

/dev-stage="Adult"
 /note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 79 a 73 c 107 g 97 t
 ORIGIN

Query Match 100.0%; Score 191; DB 10; Length 356;
 Best Local Similarity 100.0%; Pred. No. 3.8e-45;
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTTTGTAGTGGACAAAACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAG 60
 DB 270 GGCTTTGTAGTGGACAAAACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAG 211
 OY 61 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCTCAAAACCTTGACCTGACTGCACG 120
 DB 210 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCTCAAAACCTTGACCTGACTGCACG 151
 OY 121 TCCCGTGCCTCAATGCTACCTGCTCCCAATTACAGTACTTCCAAAGCAAGAGAC 180
 DB 150 TCCCGTGCCTCAATGCTACCTGCTCCCAATTACAGTACTTCCAAAGCAAGAGAC 91
 OY 181 ACCAGCAATT 191
 DB 90 ACCAGCAATT 80

RESULT 2
 LOCUS AW753448 496 bp mRNA EST 28-APR-2000
 DEFINITION QV2-CT0261-261099-011-d04 CT0261 Homo sapiens CDNA, mRNA sequence.
 ACCESSION AW753448
 VERSION AW753448.1 GI:7668380
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 496)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

FEATURES
 SOURCE 1..496
 Location/Qualifiers
 High quality sequence start: 59
 High quality sequence stop: 89.
 Location/Qualifiers
 1..496
 Location/Qualifiers
 /organism="Homo sapiens"

/db_xref="taxon:9606"
 /clone_lib="CT0261"
 /dev-stage="Adult"
 /note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 104 a 116 c 128 g 148 t
 ORIGIN

Query Match 100.0%; Score 191; DB 10; Length 496;
 Best Local Similarity 100.0%; Pred. No. 3.9e-45;
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCTTTGTAGTGGACAAAACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAG 60
 DB 464 GGCTTTGTAGTGGACAAAACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAG 405
 OY 61 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCTCAAAACCTTGACCTGACTGCACG 120
 DB 404 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCTCAAAACCTTGACCTGACTGCACG 345
 OY 121 TCCCGTGCCTCAATGCTACCTGCTCCCAATTACAGTACTTCCAAAGCAAGAGAC 180
 DB 344 TCCCGTGCCTCAATGCTACCTGCTCCCAATTACAGTACTTCCAAAGCAAGAGAC 285
 OY 181 ACCAGCAATT 191
 DB 234 ACCAGCAATT 274

RESULT 3
 LOCUS AW753451 633 bp mRNA EST 28-APR-2000
 DEFINITION QV2-CT0261-261099-011-d05 CT0261 Homo sapiens CDNA, mRNA sequence.
 ACCESSION AW753451
 VERSION AW753451.1 GI:7668383
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 633)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

FEATURES
 SOURCE 1..633
 Location/Qualifiers
 High quality sequence start: 66
 High quality sequence stop: 90.
 Location/Qualifiers
 /organism="Homo sapiens"


```
source
1. 633
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
```

BASE COUNT 136 a 151 c 161 g 185 t

ORIGIN

Query Match 100.0%; Score 191; DB 10; Length 633;
Best Local Similarity 100.0%; Pred. No. 4e-45; Indels 0; Gaps 0;
Matches 191; Conservative 0; Mismatches 0;

QY 1 GGGTTGTAGTGGACAAAACACCAAAATGGCTTACCTCCAAATCCGAGCATTTGTAAG 60
|||||
Db 462 GGGTTGTAGTGGACAAAACACCAAAATGGCTTACCTCCAAATCCGAGCATTTGTAAG 403
|||||
QY 61 GTTGGCACTTGGAAATACAGTCTGCAAGCAAGCTCACAAACCTTGACCTGACTGTACG 120
|||||
Db 402 GTTGGCACTTGGAAATACAGTCTGCAAGCAAGCTCACAAACCTTGACCTGACTGTACG 343
|||||
QY 121 TCCCGTGGCTGCAATGCTACCTGCTCCCAATTACAGTACTTCCAAAGCAAGACGAC 180
|||||
Db 342 TCCCGTGGCTGCAATGCTACCTGCTCCCAATTACAGTACTTCCAAAGCAAGACGAC 283
|||||
QY 181 ACCAGCAAAATT 191
|||||
Db 282 ACCAGCAAAATT 272.

RESULT 4
LOCUS AM361534 653 bp mRNA EST 04-FEB-2000
DEFINITION QV2-CT0261-261099-011-f01 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361534
VERSION AM361534.1 GI:6866286
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 653)
HCGP <http://www.ludwig.org.br/ORESTES>.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-CT0261-261099-011-f01&t3=1999-10-26&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 37
High quality sequence stop: 652.
Location/Qualifiers
1. 653
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"

FEATURES
source

```
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
```

BASE COUNT 140 a 156 c 166 g 191 t

ORIGIN

Query Match 100.0%; Score 191; DB 10; Length 653;
Best Local Similarity 100.0%; Pred. No. 4e-45; Indels 0; Gaps 0;
Matches 191; Conservative 0; Mismatches 0;

QY 1 GGGTTGTAGTGGACAAAACACCAAAATGGCTTACCTCCAAATCCGAGCATTTGTAAG 60
|||||
Db 517 GGGTTGTAGTGGACAAAACACCAAAATGGCTTACCTCCAAATCCGAGCATTTGTAAG 458
|||||
QY 61 GTTGGCACTTGGAAATACAGTCTGCAAGCAAGCTCACAAACCTTGACCTGACTGTACG 120
|||||
Db 457 GTTGGCACTTGGAAATACAGTCTGCAAGCAAGCTCACAAACCTTGACCTGACTGTACG 398
|||||
QY 121 TCCCGTGGCTGCAATGCTACCTGCTCCCAATTACAGTACTTCCAAAGCAAGACGAC 180
|||||
Db 397 TCCCGTGGCTGCAATGCTACCTGCTCCCAATTACAGTACTTCCAAAGCAAGACGAC 338
|||||
QY 181 ACCAGCAAAATT 191
|||||
Db 337 ACCAGCAAAATT 327

RESULT 5
LOCUS AM361532 625 bp mRNA EST 04-FEB-2000
DEFINITION QV2-CT0261-261099-011-e03 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361532
VERSION AM361532.1 GI:6866286
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 625)
HCGP <http://www.ludwig.org.br/ORESTES>.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-CT0261-261099-011-e03&t3=1999-10-26&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 93
High quality sequence stop: 608.
Location/Qualifiers
1. 625
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue

FEATURES
source

BASE COUNT 172 a 154 c 150 g 149 t
 ORIGIN mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 80.3%; Score 153.4; DB 10; Length 625;
 Best Local Similarity 99.4%; Pred. No. 3e-34;
 Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTTTGTAGTGACAAAACACCAAAATGGCTACCTCCAAATCCAGGCAATGCTAG 60
 |||||||
 DB 471 GGCTTTGTAGTGACAAAACACCAAAATGGCTACCTCCAAATCCAGGCAATGCTAG 530
 |||||||
 QY 61 GTTGGCACTTGGAAATACAGTCTGCAGAGCTGCACAACTTGACCTGACTGCAG 120
 |||||||
 DB 531 GTTGGCACTTGGAAATACAGTCTGCAGAGCTGCACAACTTGACCTGACTGCAG 590
 |||||||
 QY 121 TCCCGTGCCTCAATGCTACCTGCTCCAAATTAC 155
 |||||||
 DB 591 TCCCGTGCCTCAATGCTACCTGCTCCAAATTAC 625
 |||||||

RESULT 6
 LOCUS AM361520 621 bp mRNA EST 04-FEB-2000
 DEFINITION QV2-CT0261-261099-011-e07 CT0261 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM361520
 VERSION AM361520.1 GI:6866274
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
 TITLE The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=QV2&l2=QV2-CT0261-261099-011-e07&l3=1999-10-26&l4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 88
 High quality sequence stop: 613.
 Location/Qualifiers

FEATURES
 source 1..621
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CT0261"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site.1: Sma1; Site.2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 166 a 151 c 150 g 154 t
 ORIGIN

Query Match 71.8%; Score 137.2; DB 10; Length 621;
 Best Local Similarity 94.7%; Pred. No. 1.4e-29;
 Matches 142; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GGCTTTGTAGTGACAAAACACCAAAATGGCTACCTCCAAATCCAGGCAATGCTAG 60
 |||||||
 DB 469 GGCTTTGTAGTGACAAAACACCAAAATGGCTACCTCCAAATCCAGGCAATGCTAG 528
 |||||||
 QY 61 GTTGGCACTTGGAAATACAGTCTGCAGAGCTGCACAACTTGACCTGACTGCAG 120
 |||||||
 DB 529 GTTGGCACTTGGAAATACAGTCTGCAGAGCTGCACAACTTGACCTGACTGCAG 588
 |||||||
 QY 121 TCCCGTGCCTCAATGCTACCTGCTCCAAATTAC 150
 |||||||
 DB 589 TCCCGTGCCTCAATGCTACCTGCTCCAAATTAC 618
 |||||||

RESULT 7
 LOCUS AM361521 657 bp mRNA EST 04-FEB-2000
 DEFINITION QV2-CT0261-261099-011-f03 CT0261 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM361521
 VERSION AM361521.1 GI:6866275
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
 TITLE The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=QV2&l2=QV2-CT0261-261099-011-f03&l3=1999-10-26&l4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 123
 High quality sequence stop: 656.
 Location/Qualifiers

FEATURES
 source 1..657
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CT0261"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site.1: Sma1; Site.2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 175 a 160 c 164 g 158 t
 ORIGIN

Query Match 68.8%; Score 131.4; DB 10; Length 657;
 Best Local Similarity 95.7%; Pred. No. 6.8e-28;
 Matches 155; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GGCTTTGTAGTGACAAAACACCAAAATGGCTACCTCCAAATCCAGGCAATGCTAG 60
 |||||||
 DB 517 GGCTTTGTAGTGACAAAACACCAAAATGGCTACCTCCAAATCCAGGCAATGCTAG 576
 |||||||
 QY 61 GTTGGCACTTGGAAATACAGTCTGCAGAGCTGCACAACTTGACCTGACTGCAG 120
 |||||||
 DB 577 GTTGGCACTTGGAAATACAGTCTGCAGAGCTGCACAACTTGACCTGACTGCAG 636
 |||||||


```

/dev_stage="adult
29. .2770
CDs

```



```

/codon_start=1
/protein_id="BAB25815.1"
/db_xref="GI:12842988"
/translation="MESIKSPVFLILHLLEGVLSSELQLNNNGYEGVIAIDHDPV
EDFALIOHKMVTQASPLYFEATGKRFYRNVALIPESMKAPETTRKLEFKNA
DIVASTSPGNDPEYHEHAGCEKGRIRHLPTDFAGKRLTQYQODPTFLEMNA
PRKVENYENDEKFEYLSKGRPOARCSAITGNNOYRQCGSCICVDIRVAG
LYKDCNCFVDPHONKASITRENONISYVEFCERKHNDAPDONORCNLRSTWEV
IQSEDEKQTPMTAQPPAPFSLQIGQIVCVLDRKSSMLNDLNRNNAQSRLE
LIQVEQSGSWGVTPEPSAAVQSELKOLNSGARDLILHLPTVSAGSTSGELRT
AFVYIKKYPDGESEIVLIDGENDTISCEDLKQSGAILHPTVAGPAARKEQLS
KMGGLQYSSDOYNNGLVDFAALSSGNAIAQHSIQLESRVNLQNNQMMNGSVY
VDSGKQDTLELITMTPTPIFIIMDSGVNFIIDTTRKAYLVQPGAKGVFK
YSIOASSQDTLITVTSRASKTLPITVTVPVKNKGKPPSPVTVASISQGASPIR
ASVTAALISVNGKVTLELIDNGAGADNRKDYSSRPFRAFNDAGRYSVKTMLGCV
TSDRORAPPRKRPVYIDGTEDEVRNRPPEPTSTVODKQCFRSSTSGSVATN
VPAAPAPIDLPFCOITDLKASIQGQNLVNLTWAPDEYDHGRASNTLIMSTSIDV
LRDHPNTSLQVNTGTLIPKEASSEIEFEELGNTFGDIDFALQAVKSNLSEI
SNIRASVEIPAGPPIPEDSTPCPDISINSTIPGIHLKIMKMWLGEMQVTLGLH"
2912 2917
/note="putative"
2933
/note="putative"
polyA_site
BASE COUNT      845 a      722 c      697 g      669 t
ORIGIN

```

```

Query Match      64.8%; Score 123.8; DB 12; Length 2933;
Best Local Similarity 78.0%; Pred. No. 1.2e-25;
Matches 149; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 1 GGGTTGTAGTGACAAACCAAAATGCGCTACCTCCAAATCCAGCAATGGCTAG 60
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1652 GGTATTATCTAGACACACACACACAGGCGCTACCTCCAAATCCAGCAATGGCTAG 1711
QY 61 GTTGACACTTGGAAATATGCTGCAGCAAGCTCACAACCTTGACCTGACTGTCAAG 120
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1712 GTTGCGTTTGGAAATATGCAATTCAGAGGCTCAGACACTCTGACTGTGACAC 1771
QY 121 TCCCGTCGCTCCAAATGCTACCCGCTCCAAATGACAGTCTCCAAAGCAAGCAAGAC 180
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1772 TCCCGTCGACGACAAATGCTACCTCCTATTACAGTAGACCCGGTAGTAATTAAGAAC 1831
QY 181 ACCAGCAATTT 191
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 1832 ACAGGCAATTT 1842

```

```

RESULT 13
LOCUS      AM753449      652 bp      mRNA      EST      28-APR-2000
DEFINITION QV2-CT0261-261099-011-e08 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AM753449
VERSION     AM753449.1 GI:7668381
KEYWORDS   EST.
SOURCE      human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 652)
REFERENCE   1 (bases 1 to 652)
AUTHORS    Dias Neto, E., Garcia Correia, R., Verjovsky-Almeida, S., Briones, M. R.,
            Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,
            Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,
            Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare
            M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and
            Simpson, A. J.
COMMENT     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     Contact: Simpson A.J.G.
COMMENT     Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research

```

```

FEATURES
     source          1..652
                     Location/Qualifiers
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /clone_id="CT0261"
         /dev_stage="Adult"
         /note="Organ: colon. Vector: puc18; Site.1: SmaI; Site.2:
         SmaI; A mini-library was made by cloning products derived
         from ORESTES PCR (U.S. Letters Patent application No. 196
         /716 - Ludwig Institute for Cancer Research) profiles
         into the puc 18 vector. Reverse transcription of tissue
         mRNA and cDNA amplification were performed under low
         stringency conditions."
BASE COUNT      179 a      133 c      161 g      159 t
ORIGIN

```

```

Query Match      63.2%; Score 120.8; DB 10; Length 652;
Best Local Similarity 98.4%; Pred. No. 7.8e-25;
Matches 122; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGTTGTAGTGACAAACCAAAATGCGCTACCTCCAAATCCAGCAATGGCTAG 60
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 529 GCGTTTGTAGTGACAAACCAAAATGCGCTACCTCCAAATCCAGCAATGGCTAG 588
QY 61 GTTGACACTTGGAAATATGCTGCAGCAAGCTCACAACCTTGACCTGACTGTCAAG 120
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 529 GTTGACACTTGGAAATATGCTGCAGCAAGCTCACAACCTTGACCTGACTGTCAAG 648
QY 121 TCCC 124
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 649 TCCC 652

```

```

RESULT 14
LOCUS      AM361523      561 bp      mRNA      EST      04-FEB-2000
DEFINITION QV2-CT0261-261099-011-g05 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AM361523
VERSION     AM361523.1 GI:6866277
KEYWORDS   EST.
SOURCE      human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 561)
REFERENCE   1 (bases 1 to 561)
AUTHORS    HCGP http://www.ludwig.org.br/ORESTES.
            The FAPESP/LICR Human Cancer Genome Project
            Unpublished (1999)
JOURNAL     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?rl=QV2&ct2=QV2-CT0261-
            261099-011-g05&ct3=1999-10-26&ct4=1)
            Seq primer: puc 18 forward

```


High quality sequence start: 43
High quality sequence stop: 560.
Location/Qualifiers
1. 561

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 154 a 140 c 140 g 127 t
ORIGIN

Query Match 62.7%; Score 119.8; DB 10; Length 561;
Best Local Similarity 98.4%; Pred. No. 1.5e-24;
Matches 121; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCGTTTGTACTGACAAAAACACCAAAATGGCTTCTCAATCCAGCATTTGCTAAG 60
DB 425 GCGTTTGTACTGACAAAAACACCAAAATGGCTTCTCAATCCAGCATTTGCTAAG 484
OY 61 GTTGGACCTGGAAATACAGTCTGCAGCAAGCTCACAACCTTGACCCGACTGTCAG 120
DB 485 GTTGGACCTGGAAATACAGTCTGCAGCAAGCTCACAACCTTGACCCGACTGTCAG 544
OY 121 TCC 123
DB 545 CCC 547

RESULT 15
LOCUS BF581342 906 bp mRNA EST 12-DEC-2000
DEFINITION 602100726F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4223935 5',
mRNA sequence.

ACCESSION BF581342
VERSION BF581342.1 GI:11655054
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE NIH-MGC Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM9813 row: c column: 08
High quality sequence stop: 697.
Location/Qualifiers
1. 906

FEATURES
source

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="IMAGE:4223935"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: PCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT."

Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 251 a 237 c 246 g 172 t
ORIGIN
Query Match 59.4%; Score 113.4; DB 11; Length 906;
Best Local Similarity 78.0%; Pred. No. 1.1e-22;
Matches 149; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
OY 1 GCGTTTGTACTGACAAAAACACCAAAATGGCTTCTCAATCCAGCATTTGCTAAG 60
DB 425 GCGTTTGTACTGACAAAAACACCAAAATGGCTTCTCAATCCAGCATTTGCTAAG 484
OY 61 GTTGGACCTGGAAATACAGTCTGCAGCAAGCTCACAACCTTGACCCGACTGTCAG 120
DB 485 GTTGGACCTGGAAATACAGTCTGCAGCAAGCTCACAACCTTGACCCGACTGTCAG 544
OY 121 TCC 123
DB 545 CCC 547
Search completed: April 3, 2002, 20:27:35
Job time: 40218 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:54:01 ; Search time 521.98 Seconds
(Without alignments)
82.872 Million cell updates/sec

Title: US-09-049-696-11
Perfect score: 191
Sequence: 1 GGCTTTGTAGTGACAAAAA.....AACAGAGACACCAGCAAAATT 191

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/beckfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	30.4	15.9	10380	4	US-09-077-354B-3 Sequence 3, Appl 1
2	29	15.2	1123	3	US-09-188-930-28 Sequence 28, Appl 1
3	29	15.2	1123	3	US-09-188-930-203 Sequence 203, Appl 1
4	28.6	15.0	2172	4	US-08-760-615-1 Sequence 1, Appl 1
5	28.2	14.8	3833	5	US-08-917-320-18 Sequence 18, Appl 1
6	28.2	14.8	3833	5	PCT-US95-04611A-18 Sequence 18, Appl 1
7	28.2	14.8	5931	4	US-08-783-774-1 Sequence 1, Appl 1
8	28.2	14.8	17056	4	US-09-245-041-3 Sequence 3, Appl 1
9	27.6	14.5	2619	3	US-08-337-797A-1 Sequence 1, Appl 1
10	27.6	14.5	2619	3	US-08-337-797A-3 Sequence 3, Appl 1
11	27.6	14.5	2619	3	US-09-258-523-1 Sequence 1, Appl 1
12	27.6	14.5	2619	3	US-09-258-523-3 Sequence 3, Appl 1
13	26.8	14.0	11236	1	US-07-853-913-1 Sequence 1, Appl 1
14	26.6	13.9	1920	1	US-08-186-222-1 Sequence 1, Appl 1
15	26.6	13.9	2689	2	US-08-876-546A-15 Sequence 15, Appl 1
16	26.6	13.9	2689	2	US-09-412-252-15 Sequence 15, Appl 1
17	26.6	13.9	2689	2	US-08-204-656B-9 Sequence 9, Appl 1
18	26.6	13.9	2689	2	US-08-470-702-5 Sequence 5, Appl 1
19	26.6	13.9	2689	2	US-08-467-831-5 Sequence 5, Appl 1
20	25.8	13.5	1869	3	US-09-305-381-1 Sequence 1, Appl 1
21	25.8	13.5	2374	2	US-08-466-589-5 Sequence 5, Appl 1
22	25.8	13.5	2374	2	US-08-700-636-5 Sequence 5, Appl 1
23	25.8	13.5	2374	2	US-08-467-574-5 Sequence 5, Appl 1
24	25.8	13.5	2374	4	US-09-217-345-5 Sequence 5, Appl 1
25	25.8	13.5	4403765	4	US-09-103-840A-2 Sequence 2, Appl 1
26	25.8	13.5	4411529	4	US-09-103-840A-1 Sequence 1, Appl 1
27	25.6	13.4	1058	1	US-08-238-163-1 Sequence 1, Appl 1

ALIGNMENTS

28	25.6	13.4	4615	2	US-08-674-351-3	Sequence 3, Appl 1
29	25.4	13.3	823	4	US-08-998-416-493	Sequence 493, Appl 1
30	25.4	13.3	2580	2	US-08-511-485-7	Sequence 7, Appl 1
31	25.4	13.3	2742	3	US-08-911-853-16	Sequence 16, Appl 1
32	25.4	13.3	2742	4	US-09-479-409-16	Sequence 16, Appl 1
33	25.4	13.3	7102	3	US-09-138-024-20	Sequence 20, Appl 1
34	25.4	13.3	17612	3	US-08-911-853-29	Sequence 29, Appl 1
35	25.4	13.3	17612	4	US-09-479-409-29	Sequence 29, Appl 1
36	24.8	13.0	168	1	US-08-697-329-7	Sequence 7, Appl 1
37	24.8	13.0	1398	2	US-08-896-320-2	Sequence 2, Appl 1
38	24.8	13.0	2399	2	US-08-969-106-9	Sequence 9, Appl 1
39	24.8	13.0	2730	3	US-08-344-536-1	Sequence 1, Appl 1
40	24.8	13.0	2730	3	US-08-920-562-1	Sequence 1, Appl 1
41	24.6	12.9	920	2	US-08-860-174A-3	Sequence 3, Appl 1
42	24.6	12.9	920	4	US-09-171-025-24	Sequence 24, Appl 1
43	24.6	12.9	996	2	US-08-894-922A-4	Sequence 4, Appl 1
44	24.6	12.9	999	2	US-08-860-174A-6	Sequence 6, Appl 1
45	24.6	12.9	999	4	US-09-171-025-25	Sequence 25, Appl 1

RESULT: 1
US-09-077-354B-3
Sequence 3, Application US/09077354B
Patent No. 6255096

GENERAL INFORMATION:

APPLICANT: HOPKINS, JOHN JOSEPH, SCOTT, HAMISH STEELE;
APPLICANT: WEBER, BINGYI, BLANCH, LIANNE; ANSON, DONALD STEWART

TITLE OF INVENTION: SYNTHETIC MAMMALIAN

SEQUENCE OF INVENTION: -N-ACETYLGLUCOSAMINIDASE AND GENETIC SEQUENCES ENCODING SA

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS: ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER

STREET: 400 GARDEN CITY PLAZA

CITY: GARDEN CITY

STATE: NEW YORK

COUNTRY: UNITED STATES

ZIP: 11530

COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/077, 354B

FILING DATE: 22-APRIL-1999

PRIOR APPLICATION DATA: APPLICATION NUMBER: PCT/US96/00747

FILING DATE: 22-NOV-1996

ATTORNEY/AGENT INFORMATION: NAME: POKALSKY, ANN R.

REGISTRATION NUMBER: 34,697

REFERENCE/DOCKET NUMBER: 12416

TELEPHONE: 516 742 4343

TELEFAX: 516 742 4366

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 10380 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORGANISM: Homo sapiens

POSITION IN GENOME:

CHROMOSOME/SEGMENT: Chromosome 17

FEATURE:

NAME/KEY: exon 1

LOCATION: 990..1372

FEATURE:

NAME/KEY: exon 2
LOCATION: 2115..2262
FEATURE:
NAME/KEY: exon 3
LOCATION: 3056..3202
FEATURE:
NAME/KEY: exon 4
LOCATION: 3387..3472
FEATURE:
NAME/KEY: exon 5
LOCATION: 5667..5923
FEATURE:
NAME/KEY: exon 6
LOCATION: 7745..8955
US-09-077-354B-3

Query Match 15.9%; Score 30.4; DB 4; Length 10380;
Best Local Similarity 54.5%; Pred. No. 0.39;
Matches 61; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 51 CATTGCTAAGGTGCGCTGGAATATACAGTCTGCAAGCAAGCTCAAAACCTTGACCT 110
DB 5823 CATCTATGGGCGCGACACTTTCATAGATGACGACACCTCTCTAGAGCCCTCTACCT 5882
QY 111 GACTGTCAGCTCCGCGCTGCAATGCTACCTGCTCCATTAAGTACT 162
DB 5883 TGCCGCGACGACGACGCTCTATGAGGCAATGCTGAGTACTGCT 5934

RESULT 2
US-09-188-930-28
Sequence 28, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188.930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 28
LENGTH: 1123
TYPE: DNA
ORGANISM: Rat
US-09-188-930-28

Query Match 15.2%; Score 29; DB 3; Length 1123;
Best Local Similarity 63.8%; Pred. No. 0.41;
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 16 AAAACACCAAAATGGCTACCTCAATCCAGGCAATGCTAAGTGGCACTTGAAA 75
DB 767 aaccaagggaataatcagataccagcaaccacgagtgctgaagtgccaaaggaga 826
QY 76 TACAGCTG 84
DB 827 tgaagctg 835

RESULT 3
US-09-188-930-203
Sequence 203, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188.930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 203
LENGTH: 1123
TYPE: DNA
ORGANISM: Rat
US-09-188-930-203

Query Match 15.2%; Score 29; DB 3; Length 1123;
Best Local Similarity 63.8%; Pred. No. 0.41;
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 16 AAAACACCAAAATGGCTACCTCAATCCAGGCAATGCTAAGTGGCACTTGAAA 75
DB 767 aaccaagggaataatcagataccagcaaccacgagtgctgaagtgccaaaggaga 826
QY 76 TACAGCTG 84
DB 827 tgaagctg 835

RESULT 4
US-08-760-615-1
Sequence 1, Application US/08760615
Patent No. 6200959
GENERAL INFORMATION:
APPLICANT: Haynes, Joel R
APPLICANT: Schmaljohn, Connie S
APPLICANT: Fuller, Deborah L
APPLICANT: Schmaljohn, Alan
APPLICANT: Jahrling, Peter B
TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pluckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760.615
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 110229.91241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Ebola virus
STRAIN: Zaire
FEATURE:
NAME/KEY: CDS
LOCATION: 142..2172
OTHER INFORMATION: /product= "glycoprotein"
US-08-760-615-1

Query Match 15.0%; Score 28.6; DB 4; Length 2172;
Best Local Similarity 48.0%; Pred. No. 0.81;
Matches 82; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

OY 12 GGACAAACACCAATATGCGCTACCTCCAAATCCCGGCGATGCTAAGTTGGCACTTG 71
DB 1293 GGACAAACACCACTATATACACCCGCTGTATTAACCTTGACATCTCGAGCAACTCACT 1352
OY 72 GAATACAGTCTGCAAGCAAGCTCACAACCTTGACCCCTGACTGCTCCGCTGCGTC 131
DB 1353 TGAAACAATCACCAGGACAGACACAAGACAGACAGCCTCGAGACACTCCTGCGCAC 1412
OY 132 CAATGCTACCTGCTCCCAATTACAGTACTTCCAAAAGCAAGACAGAC 182
DB 1413 GACCGACGCGGACCCCAAAAGACAGAACACCAACACGACAGACAGAC 1463

RESULT 5
US-08-917-320-18
Sequence 18, Application US/08917320
Patent No. 5824508

GENERAL INFORMATION:
APPLICANT: Spaete, Richard and Jackman, Winthrop, T.
TITLE OF INVENTION: No. 5824508 Splicing Variants of gp350/220
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
STREET: 5 Palo Alto Square
CITY: Palo Alto
STATE: California
COUNTRY: USA

ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08/917,320
CLASSIFICATION: 435
FILING DATE: 25-AUG-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/229,291
FILING DATE: April 18, 1994

ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseri
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: AVIR-003/000US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-843-5163
TELEFAX: 415-857-0663
TELEX: 380816 CooleyPA

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
LENGTH: 3833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1014..3734

US-08-917-320-18

Query Match 14.8%; Score 28.2; DB 1; Length 3833;
Best Local Similarity 59.3%; Pred. No. 1.5;
Matches 48; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 106 ACCCTGAGTGTACGCTCCGCTGCTCAATGCTACCTGCTCCCAATTACAGTACTTCC 165
DB 2563 ACCCAGAGTGTACGCTCCCAATGCTACCTGCTCCCAATGCTACCTGCTCCCAATGCTACCT 2642

OY 106 AAAAGCAACAGACACACAC 186
DB 2643 CCAACCCCAATGCTACACAC 2663

RESULT 6
PCT-US95-04611A-18
Sequence 18, Application PC/TUS9504611A

GENERAL INFORMATION:
APPLICANT: Spaete, Richard and Jackman, Winthrop, T.
TITLE OF INVENTION: Non Splicing Variants of gp350/220
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
STREET: 5 Palo Alto Square
CITY: Palo Alto
STATE: California
COUNTRY: USA

ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04611A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,291
FILING DATE: April 18, 1994

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseri
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: AVIR-003/000US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-843-5163
TELEFAX: 415-857-0663
TELEX: 380816 CooleyPA

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
LENGTH: 3833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1014..3734

PCT-US95-04611A-18

Query Match 14.8%; Score 28.2; DB 5; Length 3833;
Best Local Similarity 59.3%; Pred. No. 1.5;
Matches 48; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 106 ACCCTGAGTGTACGCTCCGCTGCTCAATGCTACCTGCTCCCAATTACAGTACTTCC 165
DB 2563 ACCCAGAGTGTACGCTCCCAATGCTACCTGCTCCCAATGCTACCTGCTCCCAATGCTACCT 2642
OY 106 AAAAGCAACAGACACACAC 186

Page 4

RESULT 8
US-09-245-041-3
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: MOORE, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT

```
SEQUENCE: CAGGACATCATCTTCC  
LENGTH: 2619 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:
```

RESULT 8
US-09-245-041-3

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/853,913
CLASSIFICATION: 435
PRIORITY DATE: 19920319
PRIORITY APPLICATION NUMBER:
FILING DATE: 22-FEB-1991
APPLICATION NUMBER: US 07/660,412
PRIOR APPLICATION DATA:
FILING DATE: 07/603,803
APPLICATION NUMBER: US 07/201,762
FILING DATE: 02-JUN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/180,548
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-4641AAAA
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-07-853-913-1

Query Match 14.0%; Score 26.8; DB 1; Length 1126;
Best local Similarity 59.0%; Pred. No. 8.5;
Matches 46; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

OY 82 CTGCAGCAAGCTCCAAACCTTGACCTGACGTGCACGTCGCCGTGGCTCAATGCTACC 141
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6538 CTTGAGGCCAAGAACCCCGACTTGGCCAGCAGCTCCCATCCACTATATCTGAGGCGTCCC 6597

OY 142 CTGCTTCATTTACAGTG 159
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DJ 6598 TGTCTCCAATGCAGAG 6615
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-08-186-222-1
Sequence 1, Application US/08186222
Patent No. 5559007

GENERAL INFORMATION:
APPLICANT: Surl, Bruno
APPLICANT: Schmitz, Albert
TITLE OF INVENTION: Bacterial Vectors
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/186,222
FILING DATE:

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:17:59 ; Search time 1321.64 Seconds
(without alignments)
123.898 Million cell updates/sec

Title: US-09-049-696-11

Perfect score: 191
Sequence: 1 GGCTTGTACTGACCAAAA.....AACAGAGACACCAAAATT 191

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_1101.*
1: /SIDSL/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	100.0	2742	22	AAH46102 Human CLCA1 coding
2	191	100.0	2745	20	AAFE81927 Human ICAC-1, nuci
3	191	100.0	2825	22	AAH46124 Human CLCA1 CDNA,
4	191	100.0	2854	22	AAH34879 Human colon cancer
5	191	100.0	2854	22	AAH81787 Human secreted pro
6	191	100.0	2867	22	AAH33285 Human colon cancer
7	191	100.0	3109	20	AAH35019 Human membrane spa
8	191	100.0	3111	22	AAZ09840 Human membrane spa
9	191	100.0	3111	22	AAI29502 C902P determined
10	123.8	64.8	486	22	AAH46121 Mouse Gob-5 CDNA h
11	123.8	64.8	2739	22	AAH46101 Mouse Gob-5 coding

12	123.8	64.8	2843	22	AAH46120
13	122.2	64.0	2931	20	AAFE81925
14	84	44.0	2616	21	AAH461335
15	84	44.0	3265	21	AAZ65095
16	84	44.0	3265	22	AAFE92092
17	84	44.0	3265	22	AAFE44241
18	74.2	38.8	1802	21	AAH33192
19	74.2	38.8	1802	22	AAH33192
20	31	16.2	2454	22	AAZ24658
21	31	16.2	2784	20	AAZ24658
22	31	16.2	2784	20	AAZ24658
23	31	16.2	3156	21	AAZ51625
24	31	16.2	3951	20	AAZ24653
25	31	16.2	3951	21	AAZ24653
26	31	16.2	8031	31	AAZ65892
27	30.4	15.9	10380	18	AAZ67164
28	29	15.2	920	21	AAZ44789
29	29	15.2	1035	22	AAZ99776
30	29	15.2	1117	21	AAZ64064
31	29	15.2	1123	21	AAZ61633
32	29	15.2	1123	21	AAZ61730
33	29	15.2	1123	22	AAZ99566
34	29	15.2	1123	22	AAZ99663
35	29	15.2	2799	21	AAZ59995
36	28.8	15.1	2158	21	AAZ42182
37	28.6	15.0	2224	21	AAZ04041
38	28.6	15.0	2224	21	AAZ51038
39	28.6	15.0	2298	21	AAZ87189
40	28.6	15.0	2298	22	AAZ76953
41	28.6	15.0	7272	20	AAH89798
42	28.6	15.0	7272	20	AAH59393
43	28.6	15.0	7285	20	AAH89797
44	28.6	15.0	7285	20	AAZ59392
45	28.2	14.8	2721	6	AAH50114

ALIGNMENTS

RESULTS 1

ID AAH46102 standard; DNA: 2742 BP.

AC AAH46102;

XX 11-SEP-2001 (first entry)

DE Human CLCA1 coding sequence; SEQ ID NO:4.

XX Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;

XX expression inhibition; antisense therapy; gene therapy;

XX chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;

XX dsf

XX Homo sapiens;

XX Key

XX CDS

XX Location/Qualifiers

XX 1.2742

XX /tag="a

XX /product="Human CLCA1"

XX /note="No stop codon given in the specification"

XX MOZ00138530-A1.

XX 11-MAY-2001.

XX 22-NOV-2000; 2000MO-JP08232.

XX 24-NOV-1999; 99UP-0333479.

XX 27-APR-2000; 2000UP-0127589.

XX (TAKE) TAKEDA CHEM IND LTD.

XX	Nakanishi A, Morita S;
PI	
XX	DR MPI: 2001-355935/37.
XX	DR P-PDSB: AABF73716.
PT	New antisense nucleotide, useful for treatment and prevention of
PT	bronchial asthma and chronic obstructive pulmonary disease -
XX	Claim 3; Page 82-84; 104pp: Japanese.
XX	
CC	The invention relates to an antisense nucleotide targeted to the mouse
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC	the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC	relates to an antibody specific for the Gob-5 protein, medical and
CC	diagnostic compositions containing the antisense nucleotide or the
CC	antibody, and methods and kits for screening for compounds which inhibit
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC	The antisense oligonucleotides and antibody are therefore useful for the
CC	treatment and prevention of bronchial asthma and chronic obstructive
CC	pulmonary disease. The present sequence represents the human CLCA1
CC	gene coding sequence.
SO	
XX	Sequence 2742 BP: 833 A; 616 C; 623 G; 670 T; 0 other:
XX	
Query Match	100.0%; Score 191; DB 22; Length 2742;
Best Local Similarity	100.0%; Pred. No. 1,4e-55;
Matches 191; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GGCCTTTAGTGGACAAAAACGACAAATGGCCTTCACCATGCCAGCATGTCTAG 60
DB	1621 ggcttcttgtagygacaaaacacaataatgctcattcccaatcccagacttgctaag 1680
OY	61 GTTGACACTTGGAAATACAGTCTGCAAGCAAGCTCACAAACCTTGACCTGTCACG 120
DB	1681 gtggcaccttggaatatcagctctgcagaagaagctcacaaccttgaccctgactgcaag 1740
OY	121 TCCCGTGGCTCCAAATGCTACCCTGGCTCCAAATTACAGTAGCTCCAAAGCAAGAGAC 180
DB	1741 tcccglygcgtccaatgctaccctgcctccattacaglygacttcaccaaacgacaagac 1800
OY	181 ACCAGCAAATTT 191
DB	1801 accagcaaat 1811
RESULT 2	
AAF81927	
ID	AAF81927 standard; cDNA; 2745 BP.
XX	
AC	AAF81927;
XX	
DT	13-JUN-2001 (first entry)
DE	
XX	Human ICACC-1 nucleotide sequence.
XX	
KW	ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
KW	interleukin 9 induced calcium activated chloride channel; IL-9;
KW	calcium activated chloride channel; anti-allergic; anti-asthmatic;
KW	anti-inflammatory; immunomodulatory; cystic fibrosis;
KW	inflammatory bowel disease; autoimmune disease; ss.
OS	
XX	Homo sapiens.
FH	
FT	Key Location/Qualifiers
FT	CDS 1..2745
FT	/*tag= a
FT	/product= "ICACC-1"
XX	/note= "IL-9 induced calcium activated chloride channel"
PN	WO9944620-A1.
XX	

```

XX 10-SEP-1999.
XX 03-MAR-1999. 99MO-US04703.
XX 03-MAR-1998. 98US-0076815.
XX (MAGA-) MAGANININ PHARM INC.
XX HOLROYD KT, Levitt RC, Maloy WL, Louahed J, McLane M;
XX NICHOLAIDES NC, Zhou Y, Dong Q;
XX WPL, 1999-550979/46.
XX P-PSDB; AAB/4824.
XX New nucleic acid encoding calcium activated chloride channel, used to
XX identify, e.g. specific modulators for treating atopic allergy -
XX
XX Claim 1; Fig 4B; 75pp: English.
XX
XX The present sequence encodes the human interleukin 9 (IL-9) induced
XX calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
XX have anti-allergic, anti-asthmatic, anti-inflammatory and
XX immunomodulatory activities. Compounds (A) that downregulate ICACC are
XX used to alleviate asthma (or more generally atopic allergy), while those
XX (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
XX inflammatory bowel disease (IBD) (or other autoimmune diseases).
XX Measuring levels of ICACC-1 can be used in the diagnosis of asthma
XX (increased levels) or IBD (reduced levels), also for monitoring
XX treatment of these conditions. The ICACC proteins can be used:
XX (i) to raise specific antibodies (Ab), useful: (a) as immunoassay
XX reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
XX (A) to disrupt bonding between ICACC and its ligands; and (iii) to
XX identify modulators and binding partners. ICACC polynucleotides can be
XX used to generate transgenic animals or recombinant cells, used to screen
XX for antagonists, also as a source of therapeutic antisense agents or
XX diagnostic probes (for quantifying mRNA expression, e.g. for
XX identification of modulators).
XX
XX Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other:
XX
Query Match 100.0%; Score 191; DB 20; Length 2745;
Best Local Similarity 100.0%; Pred. No. 1,46-55;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 GGGTTGAGTGGACAAACCAACCAAAATGGCTCCCTCCAAATCCAGGATTGCTAAG 60
DB 1621 ggccttgagtgagcgaacaaacacacaaatgagctactcccaatcccgagatgtctaag 1600
QY 1 GTTGGCACTTGGAAATGACACTGTGCAAGCAAGCTCAACAACCTTGACCTGACTGACG 120
DB 1631 gtggcaacttggaaataacatcagtcgcgaagaagctcacaaccttgaccctgacgtcagc 1740
QY 1 TCCCGTGGCTGTCACATGCTACCTCGCTCCCAATTAAGTAGTACTTCCAAACCAACCAAGAGC 180
DB 1741 tcccggtggctgctcaatgtctaccctcgctcccaattacagtgacttccaaaagacaagaagac 1800
QY 131 ACCGACAAATT 191
DB 1801 accgacaatt 1811
XX
XX RESULT 3
XX AAH46124
XX ID AAH46124 standard; cDNA; 2825 BP.
XX AC: AAH46124;
XX
XX 11-SEP-2001 (first entry)
XX Human CLCA1 CDNA, SEQ ID NO:26.
XX
XX Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;

```


KM expression inhibition; antisense therapy; gene therapy;
KM chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic
SS.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	23..2767

```
FT      /product= "Human CLCA1"
FT      /transl_except= (pos:476..478, aa:Lys)
```

PN WO200138530-A1.

PD 31-MAY-2001

PF 22-NOV-2000; 2000WO-JP08232.

PR 24-NOV-1999; 99JP-03333479.

XX
XX
XX

XX
XX

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

DR P-PSDB; AAB73716.

PT New antisense nuc

XX Example 5; Page 92-94; 104pp; Japanese.

Example 5; Page 92-94; 104pp; Japanese.

CC The invention relates to an antisense nucleotide targetted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart.
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the Gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
CC pulmonary disease. The present sequence represents human CLCA1 cDNA.

XX Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;

90

Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;

Query Match	100.0%	Score 191; DB 22	Length 2825;
Best Local Similarity	100.0%	Pred. NC.1.4e-55;	
Matches 191; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0

QY	1643	gacctttagatgcaacaaacacccaataatgacctaccatcccaatcccaagcatgtgtaag	1702
QY	1	ggctttttacttgacacacaaaacacccaataatggcctacctcccaattccgaacattgctgaag	60
Db	1643	gacctttagatgcaacaaacacccaataatgacctaccatcccaatcccaagcatgtgtaag	1702
QY	61	gttggacacttggaataatpacagtctgcacgaagcctccaaaccttgacccttgactctgcag	120
Db	1703	gttgagcacttgggaataatacagttctgcagaagaagtcacaaacttgaacctgtgctacag	1762
QY	121	tcccggtgcgtcccaatgctaccctgctcctcccaattacagtgactttccaaacgaacaaaganc	180
Db	1763	tcccggtgcgtcccaatgctaccctgctcctcccaattacagtgactttccaaacgaacaaaganc	1822
QY	181	accagcaaat	191
Db	1823	accagcaaat	1833

RESULT	4
AAH34879	
ID	AAH34879 standard; cDNA; 2854 BP
XX	
AC	AAH34879;

03 SEP-2001 (first entry)

HuHian colon cancer antigen encoding cDNA SEQ ID NO:1961

Human; colon cancer; colon cancer antigen; diagnosis; detection
colorectal carcinoma; chromosome 1; ss.

Homo sapiens

WQZ00122920-A2.

05:APR-2001

28-SEP-2000; 2000WO-US26524.

29 SEP-1999; 99US-0157137.

10-10-68

2000

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

P-1¹SDB; AAG75474.

Nucleic acids and

Claim 1; Page 3462-3463; 9803pp; English.

[illegible]

Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query/Match	100.0%	Score 191;	DB 22	Length 2854;
Best Local Similarity	100.0%	Pred. No. 1.4e-55;		
Matches 191; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	GGCTTTTACTGTGACACAAAACACCAAAATGGCTACTCTCCAATGCCAGACATTTCTTAAG	60
DQ	1	GGCTTTTACTGTGACACAAAACACCAAAATGGCTACTCTCCAATGCCAGACATTTCTTAAG	60
DQ	1655	ggctcttgtagtggaacaaaacacacaaaatggctactctccaatcccgagcatgtgtaag	1714
QY	2	GTGGGCACTTGGAAATACAGTCTGTGCAAGCAAGCTGCACAAACCTTGACCCGTGACTGTACAG	120
DQ	2	GTGGGCACTTGGAAATACAGTCTGTGCAAGCAAGCTGCACAAACCTTGACCCGTGACTGTACAG	120
QY	1715	gttgacacttggaataacagctctggaagaagctacaacacttgagccctgactgtcaag	1774
DQ	1715	gttgacacttggaataacagctctggaagaagctacaacacttgagccctgactgtcaag	1774
QY	121	TCCCGTACGTCCCAATGTACCTCGCTCCCAATTAACGTACTTCCAAAAGCAACAAGGAC	180
DQ	121	TCCCGTACGTCCCAATGTACCTCGCTCCCAATTAACGTACTTCCAAAAGCAACAAGGAC	180
QY	1775	tcccggtgcgtcaaatgtctacccgtccctccaattacagtgacttccaaacagacaagaagac	1834
DQ	1775	tcccggtgcgtcaaatgtctacccgtccctccaattacagtgacttccaaacagacaagaagac	1834
QY	181	ACCAGCAAAATT	191
DQ	181		191
QY	1835	accagcaaat	1845
DQ	1835	accagcaaat	1845

XX	RESULT	5
XX	ID	AAAF81787
AC		AAAF81787 standard; cDNA: 2854 BP.
XX		AAAF81787;
DT		12-JUN-2001 (first entry)
XX		Human secreted protein gene 1 SEQ ID NO:11.
DE		
XX		
KW		Human: secreted protein; diagnosis; immunomodulatory; antisclerotic;
KW		dermatological; immunosuppressive; antihistaminic; anti-HIV;
KW		ophthalmological; cyclostatic; cardiac; vascular; anti-angiogenic;
KW		immunostimulant; neuroprotectant; neurotropic; anticonvulsant; vaccine;
KW		antiallergic; antiparkinsonian; antimicrobial; vulnerary; gene therapy;
KW		immune disorder; hyperproliferative disorder; cardiovascular disease;
KW		cancer; angiogenic disorder; neurological disorder; infectious disease;
KW		wound healing; regeneration; chemotaxis; chromosome 1; ss.
XX		
OS		Homo sapiens.
PN		WO200112775-A2.
XX		
PD		22-FEB-2001.
XX		
PF		16-AUG-2000; 2000WO-US22325.
XX		
PR		17-AUG-1999; 99US-0149182.
XX		
PA		(HUMA-) HUMAN GENOME SCI INC.
XX		
PI		Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;
PI		Birch CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;
XX		
DR		WPI: 2001-147550/15.
XX		P-SDB; AAF81787.
XX		
PT		Nucleic acids encoding 25 human secreted polypeptides, useful for
PT		preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT		disease and diabetic retinopathy -
XX		
PS		Claim 1; Page 441; 485pp; English.
XX		
XX		AAAF81787 to AAAF81817 encode the human secreted proteins given in AAF81787
CC		to AAF81787. Human secreted proteins can have activities based on the
CC		tissues and cells they are expressed in. Example of activities include:
CC		immunomodulatory; antisclerotic; dermatological; immunosuppressive;
CC		antihistaminic; anti-HIV; immunostimulant; cyclostatic; cardiac;
CC		vascular; anti-angiogenic; ophthalmological; neuroprotectant; neurotropic;
CC		anticonvulsant; antiallergic; antiparkinsonian; antimicrobial; and
CC		vulnerary. Human secreted proteins can be used in gene therapy and
CC		vaccine. Human secreted protein nucleotide sequences (NMI) and proteins
CC		(PEP) may be used in the prevention, diagnosis and treatment of diseases
CC		associated with inappropriate polypeptide expression. For example, NMI
CC		and PEP may be used to treat disorders or deletions in a patient's genome
CC		expression by rectifying mutation of polypeptides. Disorders that
CC		to supplement the patient's own production of polypeptides. Disorders that
CC		may be prevented, diagnosed and/or treated include immune disorders,
CC		hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
CC		angiogenic disorders, neurological disorders, infectious diseases and/or
CC		for promoting wound healing, regeneration and/or chemotaxis. AAF81787 to
CC		AAAF81786 and AAF81787 represent sequences used in the exemplification of
CC		the present invention.
XX		
XX		Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match	100.0%;	Score 191;	DB 22;	Length 2854;
Best Local Similarity	100.0%;	Pred. No. 1.4e-55;		
Matches 191;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GGCTTGTAGTGACGACAAAACACCAAAATGCGCTTACTCTCCAAATTCACAGGCAATGCTAAG	60
Db	1655	ggcttttagttagtgcacaaacacacaaatggccttacctccaatccagagcatgttaag	1714
QY	61	GTTCGGACCTGGAAATATACAGTCTGTCAGACGACAGCTCACAAACCTTGACCTGACTGTACG	120
Db	1715	gttggcacttggaaatacagtcagtcgcaagcagcttcacaaaccttgaccctgactgttaag	1774
QY	121	TCCCGTGGCGTCCAAATGCTACCCGCTCCCAATATACAGTGTGACTTCGCAAAACGACCAAGAC	180
Db	1775	tcccggtggcttgcacatgtcaccctgcctccaatctacagtgacttccaaagaaagagac	1834
QY	131	ACGAGCAATTT 191	
Db	1835	accagcaatt 1845	
RESULT	6		
AAH33285			
ID	AAH33285	standard; cDNA; 2867 BP.	
XX	AAH33285;		
XX	03-SEP-2001	(first entry)	
DI			
XX	Human colon cancer antigen encoding cDNA seq ID NO:341.		
XX	Human; colon cancer; colon cancer antigen; diagnosis; detection;		
KM	colorectal carcinoma; chromosome 1; ss.		
XX	Homo sapiens;		
OS	WC200122920-A2.		
XX	05-APR-2001.		
PD			
XX	28-SEP-2000; 2000OMO-US26524.		
XX	29-SEP-1999; 99US-0157137.		
PR	03-NOV-1999; 99US-0163280.		
FA	(HUMA-) HUMAN GENOME SCI INC.		
XX	Ruden SM, Barash SC, Birse CE, Rosen CA;		
PI	WPI; 2001-235357/24.		
XX	P-OSDB; AAG73854.		
DR			
XX	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,		
XX	useful for preventing, diagnosing and/or treating colorectal cancers -		
PT	Claim 1; Page 2452-2453; 9803pp; English.		
PS	AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon		
XX	cancer-associated nucleic acid molecules (N) and proteins (P), where		
CC	the proteins are collectively known as colon cancer antigens. The colon		
CC	cancer antigens have cytostatic activity and can be used in gene		
CC	therapy and vaccine production. N and P may be used in the prevention,		
CC	diagnosis and treatment of diseases associated with inappropriate P		
CC	expression. For example, N and P may be used to treat disorders		
CC	associated with decreased expression by rectifying mutations or deletions		
CC	in a patient's genome that affect the activity of P by expressing		
CC	inactive proteins or to supplement the patients own production of P.		
CC	Additionally, N may be used to produce the colon cancer-associated P,		
CC	by inserting the nucleic acids into a host cell and culturing the cell		
CC	to express the proteins. N and P can be used in the prevention, diagnosis		
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204		
CC	and AAG77789 represent sequences used in the exemplification of the		
CC	present invention.		
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were		
CC	missing at time of publication, meaning no sequences are present for		
CC	SEQ ID NO:1027 to 1052, 7921 and 7922.		

XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;

Query Match 100.0%; Score 191; DB 22; Length 2867;
Best Local Similarity 100.0%; Pred. No. 1.4e-55;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTTGTAGTGGACAAAACACCAAAATGGCTTACTCCAAATCCAGGATTTGCTAAG 60
|||||
DB 1659 ggccttgtagtgacaaaacacccaaaatgacctaccatcccaaggcattgctaa 1718
OY 61 GTTGACACTTGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGACCTGCTCAGC 120
|||||
DB 1719 gtggcacttggaaatacagctcgcgaagcaagctcacaaaccttgaccctgactgca 1778
OY 121 TCCCGTGGCGTCCAAATGCTACCTGCTCCAAATTAAGTACTGCTTCCAAAAGCAAGAGAC 180
|||||
DB 1779 tcccgctgctcaatgctaccctgctcccaattacagtgacttccaaaacgaacaagagac 1838
OY 181 ACCAGCAAAAT 191
|||||
DB 1839 accagcaaat 1849

RESULT 7
AAH35019
ID AAH35019 standard; cDNA; 3109 BP.

XX AAH35019;
AC
XX
XX 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:2101.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX P-PSDB; AAG75614.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 1; Page 3587-3588; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell

CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC are AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match 100.0%; Score 191; DB 22; Length 3109;
Best Local Similarity 100.0%; Pred. No. 1.4e-55;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTTTGTAGTGGACAAAACACCAAAATGGCTTACTCCAAATCCAGGATTTGCTAAG 60
|||||
DB 1508 ggccttgtagtgacaaaacacccaaaatgacctaccatcccaaggcattgctaa 1567
OY 61 GTTGACACTTGGAAATACAGTCTGCAAGCAAGCTCACAACCTTGACCTGCTCAGC 120
|||||
DB 1568 gtggcacttggaaatacagctcgcgaagcaagctcacaaaccttgaccctgactgca 1627
OY 121 TCCCGTGGCGTCCAAATGCTACCTGCTCCAAATTAAGTACTGCTTCCAAAAGCAAGAGAC 180
|||||
DB 1638 tcccgctgctcaatgctaccctgctcccaattacagtgacttccaaaacgaacaagagac 1687
OY 181 ACCAGCAAAAT 191
|||||
DB 1688 accagcaaat 1698

RESULT 8
AAZ09840
ID AAZ09840 standard; cDNA; 3111 BP.

XX AAZ09840;

XX 26-NOV-1999 (first entry)

DE Human membrane spanning protein MSP-5 cDNA fragment 2.

XX Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;
KW neoplastic disorder; immunological disorder; reproductive disorder;
KW MSP-5; ds.

XX Homo sapiens.

XX WO946380-A2.

XX 16-SEP-1999.

XX 03-MAR-1999; 99WO-US05073.

XX 13-MAR-1998; 98US-0039064.

XX (TECY-) INCYTE PHARM INC.

XX Tang YT, Bandman O, Lai P, Hillman JL, Yue H, Corley NC;

XX Giegler KJ, Kaser MR, Baughn MR, Shah P;

XX WPI; 1999-551409/46.

XX P-PSDB; AAY33298.

XX New human membrane spanning proteins used to, e.g. prevent and treat
PT neoplastic disorders -

XX Example 1; Page 80-81; 81pp; English.

XX This invention describes novel human membrane spanning proteins (MSPs),
CC and the polynucleotides encoding them. The products of the invention are
CC used to diagnose, prevent and treat neoplastic, immunological and
CC reproductive disorders. This sequence encodes a human membrane spanning

CC protein MSP-5 fragment.
XX Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match 100.0%; Score 191; DB 20; Length 3111;
Best Local Similarity 100.0%; Pred. NO. 1.4e-55;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCCTTGTAGTGACAAAACACCAAAATGGCTCTCAATCCAGCATTCCTAAG 60
DB 1654 ggccttgtatgtagtgcacaaacaccccaaatgacctcaatccagcatgctaag 1713
OY 61 GTTGGCACTTGGAAATATACAGTCTGCAGCAGCTCACAACCTTACCTCTGCTCAG 120
DB 1714 gtggcacttggaataatcagctctgcaagaagctcacaaccttgacctgactgcaag 1773
OY 121 TCCCGTGGCTCCAAATGCTACCTGCTCCAAATTCAGTCTCCAAACGAAGAAGCAG 180
DB 1774 tcccgctgctcgaatgctacccctgctcccaattacagtgacttccaaagcaagaagac 1833
OY 181 ACCAGCAATTT 191
DB 1834 accagcaaat 1844

RESULT 9
AAI29502
ID AAI29502 standard; cDNA; 3111 BP.

XX AAI29502;

DT 12-OCT-2001 (first entry)

DE C902P determined cDNA sequence.

XX Human; immunotherapy; diagnosis; colon cancer; colon tumour;

KW Immunogenic; gene therapy; vaccine; colonic cancer; ss.

XX Homo sapiens.

PN WO200149716-A2.

PD 12-JUL-2001.

PF 29-DEC-2000; 2000WO-US35596.

XX 30-DEC-1999; 99US-0476296.

PR 10-JAN-2000; 2000US-0480321.

PR 15-FEB-2000; 2000US-0504629.

PR 06-MAR-2000; 2000US-0519444.

PR 19-MAY-2000; 2000US-0575251.

PR 29-JUN-2000; 2000US-0609448.

PR 28-AUG-2000; 2000US-0649811.

XX (CORI-) CORIXA CORP.

XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;

PI King GE, Wang T, Jiang Y;

XX WPI; 2001-441847/47.

XX Colon tumor associated proteins and nucleic acids useful for the

PT prevention, diagnosis and treatment of colonic cancer -

XX Claim 2; Page 425-426; 472pp; English.

XX The present invention describes colon tumour associated proteins (I) and

CC the polynucleotides (II) that encode them. (I) have cytostatic activity.

CC (I) and (II) can be used in gene therapy and vaccine production. (I) and

CC (II) may be used in the prevention, diagnosis and treatment of diseases

CC associated with inappropriate colon tumour associated protein (TCAP)

CC expression, such as colonic cancer. For example, (I) and (II) may be

used to treat disorders associated with decreased expression by
rectifying mutations or deletions in a patient's genome that affect the
activity of TCAPs by expressing inactive proteins or to supplement the
patients own production of them. Additionally, (II) may be used to
produce the TCAP proteins, by inserting the nucleic acids into a host
cell culturing the cell to express the protein. (II) and its
complementary sequences may also be used as DNA probes in diagnostic
polymerase chain reaction (PCR) and hybridisation assays to detect and
quantitate the presence of similar nucleic acids in samples, and
therefore which patients may be in need of restorative therapy. (I) may
also be used as antigens in the production of antibodies against TCAPs
and in assays to identify modulators of TCAP expression and activity.
Anti-(I) antibodies and antagonists may also be used to down regulate
TCAP expression and activity. The anti-(I) antibodies may also be used
as diagnostic agents for detecting the presence of TCAPs in samples
(e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
and AAI24494 to AAI24523 represent nucleotide and amino acid sequences
given in the exemplification of the present invention.

Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;

Query Match 100.0%; Score 191; DB 22; Length 3111;
Best Local Similarity 100.0%; Pred. NO. 1.5e-55;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCCTTGTAGTGACAAAACACCAAAATGGCTCTCAATCCAGCATTCCTAAG 60
DB 1972 ggccttgtatgtagtgcacaaacaccccaaatgacctcaatccagcatgctaag 2031

OY 61 GTTGGCACTTGGAAATATACAGTCTGCAGCAGCTCACAACCTTACCTCTGCTCAG 120

DB 2032 gtggcacttggaataatcagctctgcaagaagctcacaaccttgacctgactgcaag 2091

OY 121 TCCCGTGGCTCCAAATGCTACCTGCTCCAAATTCAGTCTCCAAACGAAGAAGCAG 180

DB 2092 tcccgctgctcgaatgctacccctgctcccaattacagtgacttccaaagcaagaagac 2151

OY 181 ACCAGCAATTT 191

DB 2152 accagcaaat 2162

RESULT 10

FAH46121
ID FAH46121 standard; cDNA; 486 BP.

XX FAH46121;

AC 11-SEP-2001 (first entry)

DT Mouse Gob-5 cDNA hybridisation probe, SEQ ID NO:23.

XX Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;

KW expression inhibition; antisense therapy; gene therapy; bronchial asthma;

KW chronic obstructive pulmonary disease; antiasthmatic;

KM hybridisation probe; ss.

XX Mus sp.

XX WO200138530-A1.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-JP08232.

XX 26-NOV-1999; 99JP-0333479.

XX 27-APR-2000; 2000JP-0127589.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Nakanishi A, Morita S;

PI

31-MAY-2001.

22-NOV-2000; 2000MO-JP08232.

24-NOV-1999; 99JP-0333479.

27-APR-2000; 2000JP-0127589.

(TAKE) TAKEDA CHEM IND LTD.

Nakanishi A, Morita S;

WPI: 2001-355935/37.

P-PSDB; AAB73715.

New antisense nucleotide, useful for treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease -

Example 1; Page 89-91; 104pp; Japanese.

The invention relates to an antisense nucleotide targeted to the mouse Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart, the CLCA1 gene (coding sequence shown in AAH46102). The invention also relates to an antibody specific for the Gob-5 protein, medical and diagnostic compositions containing the antisense nucleotide or the antibody, and methods and kits for screening for compounds which inhibit the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells. The antisense oligonucleotides and antibody are therefore useful for the treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease. The present sequence represents mouse Gob-5 cDNA.

Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;

Query Match 64.8%; Score 123.8; DB 22; Length 2843; Best Local Similarity 78.0%; Pred. No. 1.5e-32; Matches 149; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

1 GGGTTGTAGTGGACAAAACACCAAAATGGCTTCCAAATCCAGCATTTGCTAG 60
 1638 ggtttatctagacacacactaagtgctctaccctcgaagtcacagcagctaa 1697
 61 GTTGGCACTTGGAAATACAGTCTGCAAGCAAGCTCAGAACTTGACCTGCTACG 120
 1698 gtggcttttggaatacagcattcaagcagctcacaagctcactcctgactgtcacc 1757
 121 TCCCGTGGCTGCAATGCTGCTCCTCAATACAGTGAAGTCTCAAAAGCAAGAGAC 180
 1758 tcccggtgcagcaagtgctacactgtcctctattacagtgaccccggtagtgaataaac 1817
 181 ACCAGCAAAAT 191
 1818 acagggaaatt 1828

RESULT 13
 AAF81925
 ID AAF81925 standard; cDNA; 2931 BP.
 AC AAF81925;
 XX
 DT 13-JUN-2001 (first entry)
 XX
 DE Murine ICACC-1 nucleotide sequence.
 XX
 ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder; Interleukin 9 induced calcium activated chloride channel; IL-9; calcium activated chloride channel; anti-allergic; anti-asthmatic; anti-inflammatory; immunomodulatory; cystic fibrosis; inflammatory bowel disease; autoimmune disease; ss.
 OS Mus sp.
 XX
 FH Key Location/Qualifiers

8.2749
 /*tag= a
 /product= "ICACC-1"
 /note= "IL-9 induced calcium activated chloride channel"

03-MAR-1999; 99MO-US04703.
 03-MAR-1998; 98US-0076815.
 (MAGA-) MAGANIN PHARM INC.
 Holroyd KJ, Levitt RC, Maloy WL, Louheh J, McLane M;
 Nigolides NC, Zhou Y, Dong Q;
 WPI: 1999-550979/46.
 P-PSDB; AAB74822.

New nucleic acid encoding calcium activated chloride channel, used to identify, e.g. specific modulators for treating atopic allergy -

Claim 2; Fig 2; 75pp; English.

The present sequence encodes the murine interleukin 9 (IL-9) induced calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins have anti-allergic, anti-asthmatic, anti-inflammatory and immunomodulatory activities. Compounds (A) that downregulate ICACC are used to alleviate asthma (or more generally atopic allergy), while those (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and inflammatory bowel disease (IBD) (or other autoimmune diseases). Measuring levels of ICACC-1 can be used in the diagnosis of asthma (increased levels) or IBD (reduced levels), also for monitoring treatment of these conditions. The ICACC proteins can be used: (1) to raise specific antibodies (Ab), useful: (a) as immunoassay reagents, and (b) as therapeutic (A); (11) as fragments or mutants, as (A) to disrupt bonding between ICACC and its ligands; and (11) to identify modulators and binding partners. ICACC polynucleotides can be used to generate transgenic animals or recombinant cells, used to screen for antagonists, also as a source of therapeutic antisense agents or diagnostic probes (for quantifying mRNA expression, e.g. for identification of modulators).

Sequence 2931 BP; 867 A; 719 C; 684 G; 661 T; 0 other;

Query Match 64.0%; Score 122.2; DB 20; Length 2931; Best Local Similarity 77.5%; Pred. No. 5.5e-32; Matches 148; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

1 GGGTTGTAGTGGACAAAACACCAAAATGGCTTCCAAATCCAGCATTTGCTAG 60
 1631 ggtttatctagacacacactaagtgctctaccctcgaagtcacagcagctaa 1690
 61 GTTGGCACTTGGAAATACAGTCTGCAAGCAAGCTCAGAACTTGACCTGCTACG 120
 1691 gtggcttttggaatacagcattcaagcagctcacaagctcactcctgactgtcacc 1750
 121 TCCCGTGGCTGCAATGCTGCTCCTCAATACAGTGAAGTCTCAAAAGCAAGAGAC 180
 1751 tcccggtgcagcaagtgctacactgtcctctattacagtgaccccggtagtgaataaac 1810
 181 ACCAGCAAAAT 191
 1821 acagggaaatt 1821

RESULT 14
 AAB64335
 ID AAB64335 standard; DNA; 2616 BP.
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:02:04 ; Search time 22700.8 Seconds
(without alignments)
138,804 Million cell updates/sec

Title: US-09-049-696-11

Perfect score: 191
Sequence: 1 GGCTTGTAGTGACAAAAA.....AACAGACACCAGCAATT 191

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rnd:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rnd:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	191	100.0	2022	9	AK024970	AK024970 Homo sapi
2	191	100.0	2826	9	AF127036	AF127036 Homo sapi
3	191	100.0	3311	6	AX193489	AX193489 Sequence
4	191	100.0	3311	9	AF039400	AF039400 Homo sapi
5	135.8	71.1	35278	9	AF039401	AF039401 Homo sapi
6	135.8	71.1	113764	2	HS1651E10	AL122002 Human DNA
7	135.8	71.1	164891	2	AL358950	AL358950 Homo sapi
8	124.4	65.1	3079	4	AF095584	AF095584 Sus scrofa
9	123.8	64.8	2937	10	AB017156	AB017156 Mus muscu
10	84	44.0	1895	9	AK000138	AK000138 Homo sapi
11	84	44.0	3204	9	AF127035	AF127035 Homo sapi
12	84	44.0	3221	9	AK000072	AK000072 Homo sapi
13	84	44.0	3265	6	AX092338	AX092338 Sequence
14	56.6	29.6	140718	2	AL356270	AL356270 Homo sapi
15	35.8	18.7	2820	4	AF001263	AF001263 Bos tauru
16	35.8	18.7	3288	4	AF001262	AF001262 Bos tauru
17	35.8	19.7	3317	4	AF001261	AF001261 Bos tauru
18	35	18.3	153299	2	AC073593	AC073593 Homo sapi
19	35	18.3	183928	2	AC020773	AC020773 Homo sapi
20	34.8	18.2	122272	9	AL359845	AL359845 Human DNA
21	32.6	17.1	2832	9	AF127980	AF127980 Homo sapi
22	32.6	17.1	4077	6	AX054697	AX054697 Sequence
23	32.2	16.9	124775	2	AC010653	AC010653 Homo sapi
24	32.2	16.9	125768	2	AC026451	AC026451 Homo sapi
25	32	16.8	3032	10	AF047838	AF047838 Mus muscu
26	32	16.8	3471	10	AF052746	AF052746 Mus muscu
27	31.8	16.6	188361	9	AC087880	AC087880 Homo sapi
28	31.8	16.6	253622	2	AC084880	AC084880 Homo sapi
29	31.6	16.5	66348	2	AC025016	AC025016 Homo sapi
30	31.6	15.5	107642	10	AF125313	AF125313 Mus muscu
31	31.6	16.5	146478	2	AC032036	AC032036 Homo sapi
32	31.6	16.5	149252	9	AP002001	AP002001 Homo sapi
33	31.6	16.5	162529	2	AP001985	AP001985 Homo sapi
34	31.6	16.5	165729	2	AC009642	AC009642 Homo sapi
35	31.6	16.5	181459	2	AC053477	AC053477 Homo sapi
36	31.4	16.4	131234	9	HS172K2	284814 Human DNA s
37	31.2	15.3	2435	9	BC004874	BC004874 Homo sapi
38	31.2	16.3	5319	9	AB051503	AB051503 Homo sapi
39	31.2	16.3	34175	9	CNS00YVB	AL096805 Homo sapi
40	31.2	16.3	109880	2	AF181895	AF181895 Homo sapi
41	31.2	16.3	110736	2	AL162741	AL162741 Homo sapi
42	31.2	16.3	118555	2	AL139287	AL139287 Homo sapi
43	31.2	16.3	185178	2	AC013561	AC013561 Homo sapi
44	31.2	16.3	216045	2	AF128834	AF128834 Homo sapi
45	31	16.2	2048	3	M0SGAAR	L14606 Aedes aegypt

ALIGNMENTS

RESULT 1
AK024970
LOCUS
DEFINITION Homo sapiens cDNA: FLJ21317 fls, clone COL02275, highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CACCL1) mRNA.
ACCESSION AK024970
VERSION AK024970.1 GI:10437397
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_1lb:COL clone:COL02275.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Odayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
MEDO human cDNA sequencing project
Unpublished (2000)
TITLE JOURNAL
REFERENCE 2 (bases 1 to 2022)

AUTHORS Sugano, S., Suzuki, Y., Ota, T., Ohashi, M., Nishi, T., Isogai, T., Shihahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sunko Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry for Biotechnology; cDNA library sequencing; Research Association for Biotechnology; department of construction, 5'-8'-3'-end one pass sequencing; department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES
source 1. 2022
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL02275"
/clone_11b="COL"
/tissue_type="colon"
/note="cloning vector pME18SFL3"
1. 2022
/note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA"
BASE COUNT 612 a 472 c 453 g 485 t
ORIGIN
Query Match 100.0%; Score 191; DB 9; Length 2022;
Best Local Similarity 100.0%; Pred. No. 4.9e-51;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGGTTGTAGTGCACAAAACACCAAAATGCGCTACCTCCAAATCCAGGATTCCTAAG 60
Db 834 GCGTTTGTAGTGCACAAAACACCAAAATGCGCTACCTCCAAATCCAGGATTCCTAAG 893
1 GGGTTGTAGTGCACAAAACACCAAAATGCGCTACCTCCAAATCCAGGATTCCTAAG 120
Db 61 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCAAGTCCAAACCTGAGCCGACTGTCAG 933
1 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCAAGTCCAAACCTGAGCCGACTGTCAG 933
Db 121 TCCCGTGCCTGCAATGCTACCTGCTCCATTAACAGTGAATCCAAACGAACAGAGAC 180
1 TCCCGTGCCTGCAATGCTACCTGCTCCATTAACAGTGAATCCAAACGAACAGAGAC 180
Db 954 TCCCGTGCCTGCAATGCTACCTGCTCCATTAACAGTGAATCCAAACGAACAGAGAC 1013
1 TCCCGTGCCTGCAATGCTACCTGCTCCATTAACAGTGAATCCAAACGAACAGAGAC 1013
Db 181 ACCAGCAAAATT 191
1 ACCAGCAAAATT 191
Db 1014 ACCAGCAAAATT 1024
1 ACCAGCAAAATT 1024
RESULT 2
AF127036 2826 bp mRNA PRI 10-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)
DEFINITION mRNA, complete cds.
ACCESSION AF127036
VERSION AF127036.1 GI:4585468
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2826)
AUTHORS Agnel, M., Verma, T. and Culouscou, J. M.
TITLE Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea
JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792
REFERENCE 2 (bases 1 to 2826)

AUTHORS Agnel, M. and Culouscou, J.-M.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrieres, Rueil-Malmaison 92500, France
FEATURES
source 1. 2826
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="small intestine; colon"
1. 2826
/gene="CaCC1"
5. 2749
/gene="CaCC1"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 1"
/protein_id="A025487.1"
/db_xref="GI:4585468"
/translation="MGPFKSSVFILILHLEGLSLSLQLNNNGEIVADIPRV
EDFTLILOIKDMVTOASLYLFEATGRFFKFNVALILPEFTKADYVPRKLETKNA
DVLVAESTPPGNDPEYEQMGNGEGERHILPDIACKLAIEYGPQRAVHEVAH
LRNGVPEYNDKEFTLSNGRIQAVCSAGITGVNKKCGSGCYTRCTFNKTVY
YKGCFTVLOSRTKASITMFAOHVSIVEFCEONHNEAPKOKCNLSNTEVI
RDSBDKRTTPPTQPPNPFESLIQIGQRTVCYLKSGSMAGNLANLNOGOLEFL
LQVEIGSWGMYTPDSAAHVOSLLIQINGSDPDLAKRPAASGCTSGSLPSA
FVYIRKYPIDGSEIVLLTDEGNTISGCNEKQSGALIHVALDPSAQBELBLK
MTGGLQTVADDOVONNGLIDAFGLSNGNVAQSRIQLESGLTLQNSQOMNGVIV
DSYVGKDTFLITWTTPQIILMDPSGQGGGVVDKNTKAVYLOIPIDAKGVTKY
SLOASSQTLITVTSRASNTLPIITVSTKNDTSKFPSPILVYANIRQASPIIRA
SVRLTISVNGKIVTLELLDNGAGADATKDDGYSHREFTYDNGRYSKVRALGSVN
AARNAVIPQSGALYIRGWIENDEIOMNPPRPIINDDOHOKOVCSGSEFAS
DVNPAPIPDLPEQITDLKAEIHGSLINULTATAGDDYDHTAKYIIRISTSLD
LRDFNLSLOYNTALIPKANESEVLEFEPENTENGDTLFTAIQADKVDLKEI
SNIRVSLFIPQTPPETPSPDETSAFCPIHINSTITPGIHLIKIMKWIGELQLSIA
BASE COUNT 875 a 623 c 632 g 696 t
ORIGIN
Query Match 100.0%; Score 191; DB 9; Length 2826;
Best Local Similarity 100.0%; Pred. No. 5e-51;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GCGTTTGTAGTGCACAAAACACCAAAATGCGCTACCTCCAAATCCAGGATTCCTAAG 60
Db 1625 GCGTTTGTAGTGCACAAAACACCAAAATGCGCTACCTCCAAATCCAGGATTCCTAAG 1684
1 GCGTTTGTAGTGCACAAAACACCAAAATGCGCTACCTCCAAATCCAGGATTCCTAAG 120
Db 51 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCAAGTCCAAACCTGAGCCGACTGTCAG 1744
1 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCAAGTCCAAACCTGAGCCGACTGTCAG 1744
Db 1695 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCAAGTCCAAACCTGAGCCGACTGTCAG 1804
1 GTTGGCACTTGGAAATACAGTCTGCAGCAAGCAAGTCCAAACCTGAGCCGACTGTCAG 1804
Db 1745 TCCCGTGCCTGCAATGCTACCTGCTCCATTAACAGTGAATCCAAACGAACAGAGAC 1804
1 TCCCGTGCCTGCAATGCTACCTGCTCCATTAACAGTGAATCCAAACGAACAGAGAC 1804
Db 181 ACCAGCAAAATT 191
1 ACCAGCAAAATT 191
Db 1805 ACCAGCAAAATT 1815
1 ACCAGCAAAATT 1815
RESULT 3
AX193489 3311 bp DNA PAT 15-AUG-2001
LOCUS AX193489
DEFINITION Sequence 1056 from Patent WO0149716.
ACCESSION AX193489
VERSION AX193489.1 GI:15211440
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3311)
AUTHORS Xu, J., Lodes, M. J., Secrist, H., Benson, D. R., Mesgher, M. J.,

TITLE Stolk, J.A., King, G.E., Wang, T. and Jiang, Y.
Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
Patent: WO 0149716-A 1056 12-JUL-2001,
CORIXA CORPORATION (US)

FEATURES
source location/Qualifiers
1..3311
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 100.0%; Score 191; DB 6; Length 3311;
Best local Similarity 100.0%; Pred. No. 5e-51;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTTGTAGTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAAG 60

Db 1972 GCGTTGTAGTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAAG 2031

QY 61 GTTGGACCTGGAAATACAGTCTGCAAGCAAGCTCAACAACTTGACCTGACTGTCACG 120

Db 2032 GTTGGACCTGGAAATACAGTCTGCAAGCAAGCTCAACAACTTGACCTGACTGTCACG 2091

QY 121 TCCCGTCGCTGCAATGCTACCTGCTCCAAATTAACAGTGAATTCACAAAGCAAGGAC 180

Db 2092 TCCCGTCGCTGCAATGCTACCTGCTCCAAATTAACAGTGAATTCACAAAGCAAGGAC 2151

QY 181 ACCAGCAAAATT 191

Db 2152 ACCAGCAAAATT 2162

RESULT 4
AF039400 3311 bp mRNA PRI 14-DEC-1998

LOCUS Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA,
complete cds.

ACCESSION AF039400
AF039400.1 GI:4009457

KEYWORDS
SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3311)
Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Gruber, A.D., Elble, R.C., Ji, H.L., Schreuer, K.D., Fuller, C.M. and
Pauli, B.U.

TITLE Genomic cloning, molecular characterization, and functional
analysis of human CLCA1, the first human member of the family of
Ca2+-activated Cl- channel proteins

JOURNAL Genomics 54 (2), 200-214 (1998)

MEDLINE 99047526

REFERENCE 2 (bases 1 to 3311)
Gruber, A.D., Elble, R. and Pauli, B.U.

AUTHORS Direct Submission

TITLE Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
Cornell University, Ithaca, NY 14853, USA

JOURNAL Location/Qualifiers

FEATURES
source 1..3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/tissue-type="small intestine"
1..3311
/gene="hCLCA1"
352..3096
/gene="hCLCA1"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"

/protein_id="AAC95428.1"

/db_xref="GI:4009458"
/translation="MGPKSVFILLHLLEGALSNSLIQLNNNGEIVVAIDNPV
EDLLIQIKDWVQASLYLFEATGKGEKREIVAILIPETWTKADYVPRKLEKNA
DYLVAESTPRGNDPEPTBOMNGCKEGERIHLTPDIAKKAIEYGPCKAYVHEAH
LRMGVDEYNDEKFTLSNGRIQAHKCSKIGIVNKKCGGSCCTKNTKVTGL
YKGECEVYQSRQTERKASIMFAOHDSIVFCTEDHNNKRNKQKCNLSTWVFI
RSEDEKFTKPTMTQPNPFPDLQIGRIIVCLVLDGSGSMATGNLNLNAGOLF
LOTVELGYSWGVTFEDSAHVOSELIQINSQSDRLAKRLPAASGGTSIGSLSA
FTVIRKRYPTDGESEIILLTDGDNITSCGFENKOSGAIHIVVALGSPAQBLELSK
MTGGLQTVASDVOVNNGLDARGALSSNGAYSQRSIOLESGLITLONNNGYIV
DSTVGKDTLFLTWTPQPIILLMPDSGKGGFVVDKNTKMAVLIQIPAIAGVTKY
SLQASQDTLTLVTSRASNATLPPIVTSKTKDYSKFSPPLVYANIRQASPTLRA
SVTALIESYNGKTVLELLDNGAGADATDDQVSRFFPTDNGRYSVKVALGCVN
AARRVYIPQOSGALVYIPGWIENDEIOMNPREDINKDDYOHKOVCFSTSGSPVAS
DYPNADIPDLPPRGOTTDKATIHGSLINLTWAPGDYDGTAKHYIIRISTSID
LQDKRESIQVNTTALIRKANSSEVFLEKPENTITFENGTDLFIQIADYKVDLSEI
SNIRVSLFIPQTPPTPSPDSTAPCPNIIHINSTIPGIIHLKIMKWIQELQLSIA

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 100.0%; Score 191; DB 9; Length 3311;
Best local Similarity 100.0%; Pred. No. 5e-51;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTTGTAGTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAAG 60

Db 1972 GCGTTGTAGTGGACAAAACACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAAG 2031

QY 61 GTTGGACCTGGAAATACAGTCTGCAAGCAAGCTCAACAACTTGACCTGACTGTCACG 120

Db 2032 GTTGGACCTGGAAATACAGTCTGCAAGCAAGCTCAACAACTTGACCTGACTGTCACG 2091

QY 121 TCCCGTCGCTGCAATGCTACCTGCTCCAAATTAACAGTGAATTCACAAAGCAAGGAC 180

Db 2092 TCCCGTCGCTGCAATGCTACCTGCTCCAAATTAACAGTGAATTCACAAAGCAAGGAC 2151

QY 181 ACCAGCAAAATT 191

Db 2152 ACCAGCAAAATT 2162

RESULT 5

AF039401 35278 bp DNA PRI 14-DEC-1998

LOCUS Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene,
complete cds.

ACCESSION AF039401
AF039401.1 GI:4009459

KEYWORDS
SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 35278)
Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Gruber, A.D., Elble, R.C., Ji, H.L., Schreuer, K.D., Fuller, C.M. and
Pauli, B.U.

TITLE Genomic cloning, molecular characterization, and functional
analysis of human CLCA1, the first human member of the family of
Ca2+-activated Cl- channel proteins

JOURNAL Genomics 54 (2), 200-214 (1998)

MEDLINE 99047526

REFERENCE 2 (bases 1 to 35278)
Gruber, A.D., Elble, R. and Pauli, B.U.

AUTHORS Direct Submission

TITLE Submitted (19-DEC-1997) Department of Pathology, College of
Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

JOURNAL Location/Qualifiers

FEATURES
source 1..35278
/organism="Homo sapiens"
/db_xref="taxon:9606"

Chromosome="1"
map="1p22-p31"
1..1617
/gene="hCLCA1"
1..33522
/gene="hCLCA1"
767..1265
repeat_region
/rpt_family="L1 transposable element"
/rpt_type=dispersed
1587..1596
/gene="hCLCA1"
/note="putative"
1618..1708
/gene="hCLCA1"
/number=1
mRNA
join(1618..1708,1962..2383,6667..6807,6979..7126,9696..9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33310)
/gene="hCLCA1"
/product="calcium-dependent chloride channel-1"
1962..2383
/gene="hCLCA1"
/number=2
CDS
join(2222..2383,6667..6807,6979..7126,9696..9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33310)
/gene="hCLCA1"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95429.1"
/db_xref="GI:4009460"
/translation="MGPKRSVFIIILHLEGAISNSIIQLNNGYEICIVADINPTPEBETLIIQIKDMVTOASLYLFPEATGRKRYFFNNVALLPETWKTADVYRPLEYKNAADVLVAESTPGNDEPYETBOMGNCGEKGRIRILTPDLFIAGKRLAEYGDGKAFFEMAHLRWGQVDEVDNDEKFIYLSNGRIQAVRCSAGITGVVKKCGGSGYTRCFNFVTGIGLIRBCEGCVIYOSRQTERKASITMFAOHVDSIYECCTEDNNHKRPNKONOKC.NRSWEVLYRSDERKVTITPTQPPNPTFSLIQIGRIYCVLIDKSGSWATNRELRNDAQDLTLRQVEDKSVGWAVTFDSAHHVQSELIQINSQSDRTLKKRLPAASGCTSISSLRBRAPFYIRKRYPTDSEIYVLTLDGEDNTISGFENVKOSGAIHTVALGPSAAQDELERSKMTGLOTFASDOVQNNGLIDAFGALSSGANGAVSOKSIOLERSGLIIONSMONNTVYKDSVSGDTEFLITVSTQPPQILIMPDSOKOGFVVDKNTKMAYLQIPIGIAKVTVTWYSIOASQDTLITVTSRASNAWLPITVYTSIKTRQNSKPPSLVYVANIROGASPLIRASVALIESYNGKTVLTLELDNGAGADAKKDDGVSRYFTYTDINGRSVSKVALGGVNAARRRITIPQSGALYIPGWTENDEYQNMPPREIYKDDVQKHQKQVSEPTSSGSEFVASDVNARRITPDLFPPGQITLDLNAIEHGSSLNLTWAPGDDTDHGTAKHTIIITISILDDLRKQFNEISLQVNTALIPKEANSSEVFLERPNITFENGTDLFIATQAVDVVDLKSELSNRARVFSLIPQTPETPSPDSTSAPCPNHININSTIPGIIHLKIMRWIGELIOLISIA"

BASE COUNT	ORIGIN
11245	a 6870 c 7230 g 9933 t
exon	/number=10 2664. .26859 /gene="hCLICAL"
exon	/number=11 27447. .27708 /gene="hCLICAL"
exon	/number=12 28766. .28936 /gene="hCLICAL"
exon	/number=13 31840. .32079 /gene="hCLICAL"
exon	/number=14 32919. .33522 /gene="hCLICAL"
exon	/number=15

Query Match	71.1%	Score 135.8	DB 9	Length 35278
Best Local Similarity	98.6%	Pred. No. 4e-33		
Matches 137	Conservative	0	Mismatches 2	Indels 0
				Gaps 0
QY	TTCTCAAGCTTGGACATTGGAAATACAGTCTGCAAGAGCTCAAAACCTTGACCTGA	112		
DQ	27439 TTGGTCAGGTGGSCAOTTGGAAATACAGTCTGCAAGAGCTCAAAACCTTGACCTGA	27498		
QY	113 CTGTCAAGTCCGCGTGCCAAATGCTACCGCTGCTCCAAATTACAGTACTTCCAAAACA	172		
DQ	27439 CTGTCAAGTCCGCGTGCCAAATGCTACCGCTGCTCCAAATTACAGTACTTCCAAAACA	27558		
CY	173 ACAAGGACACCAACAATTT 191			
DQ	27539 ACAAGGACACCAACAATTT 27577			

RESULT 6
HSJ651E10
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

HSJ651E10 113764 bp DNA PRI 27-MAY-2000
Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1
complete sequence.
AL122002
AL122002.16 GI:8247274
HTG.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 113764)
Brown, A.
Direct Submission
Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,

COMMENT

Requests: clonerequestsanger.ac.uk
On Jun 4, 2000 this sequence version replaced g1:8247030.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
EM, EMBL; SW, SWISSPROT; TR, TREMBL; WP, WORMPEP; Information
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Ch1>
RP4-651E10 is from the library RP4-651E10 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>
VECTOR: pCYPAC2

This sequence is the entire insert of clone RP4-651E10.

FEATURES

Source

Location/Qualifiers

1..113764

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/map="p22.3-31.1"

/clone="RP4-651E10"

/clone_lib="RP4-4"

BASE COUNT 34890 a 21989 c 22351 g 34534 t

ORIGIN

Query Match

Best Local Similarity 98.6%; Pred. No. 4.3e-33; Length 113764;
Matches 137; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 53 TTGCTAGGTTGGCACTTGGAAATACAGTCTGCAAGCAACCTTGACCTGA 112

Db 27149 TTGCTAGGTTGGCACTTGGAAATACAGTCTGCAAGCAACCTTGACCTGA 27208

QY 113 CTGTACAGTCCGTCGCTCAATGCTACCTGCTCAATTACAGTCTGCAAGCA 172

Db 27209 CTGTACAGTCCGTCGCTCAATGCTACCTGCTCAATTACAGTCTGCAAGCA 27268

QY 173 ACAAGGACACCAACAATT 191

Db 27269 ACAAGGACACCAACAATT 27287

RESULT 7

AL358950/c

LOCUS

AL358950 164891 bp DNA HTG 23-JAN-2001

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL358950 164891 bp DNA HTG 23-JAN-2001
Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN
PROGRESS *** 10 unordered pieces.
AL358950.4 GI:12539689
HTG: HTGS_PHASE1; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 164891)
Plumb, B.
Direct Submission
Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Jan 26, 2001 this sequence version replaced gi:9988471.
Genome Center
Center: Sanger Centre
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk
Project Information
Center project name: BA444C12
Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least Q40
Consensus quality: 162610 bases at least Q30
Consensus quality: 163328 bases at least Q20
Insert size: 163991; sum-of-contigs
Insert size: 163373; 6.6% error; agarose-ff

Quality coverage: 5.19x in Q20 bases; sum-of-contigs Quality coverage: 5.21x in Q20 bases; agarose-ff

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 34236: contig of 34236 bp in length
34237 34336: gap of 100 bp
34337 47372: contig of 13036 bp in length
47373 47472: gap of 100 bp
47473 67025: contig of 19553 bp in length
67026 67125: gap of 100 bp
67126 71239: contig of 4114 bp in length
71240 71339: gap of 100 bp
71340 83903: contig of 12564 bp in length
83904 84003: gap of 100 bp
84004 91357: contig of 7354 bp in length
91358 91457: gap of 100 bp
91458 96452: contig of 4995 bp in length
96453 96552: gap of 100 bp
96553 148189: contig of 51637 bp in length
148190 148289: gap of 100 bp
148290 159778: contig of 11489 bp in length
159779 159878: gap of 100 bp
159879 164891: contig of 5013 bp in length.

FEATURES

Source

Location/Qualifiers

1..164891

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/clone="RP11-444C12"

/clone_lib="RP4-4"

1..34236

/note="assembly:fragment:00027"

fragment_chain:1

clone_end:77

vector_side:left

34337..47372

/note="assembly:fragment:00312"

fragment_chain:1

47473..67025

/note="assembly:fragment:01923"

fragment_chain:1

67126..71239

/note="assembly:fragment:02094"

fragment_chain:1

71340..83903

/note="assembly:fragment:02213"

fragment_chain:1

84004..91357

/note="assembly:fragment:01292"

fragment_chain:1

91458..96452

/note="assembly:fragment:01391"

fragment_chain:1

96553..148189

/note="assembly:fragment:02045"

fragment_chain:1

148290..159778

/note="assembly:fragment:00427"

fragment_chain:1

159879..164891

/note="assembly:fragment:00773"

fragment_chain:1

clone_end:SP6

vector_side:right

BASE COUNT 48381 a 32179 c 32357 g 51067 t 907 others

ORIGIN

LLQIVEGSGWGTGTFPSAAYQSEIKOLNSGADNDLIKHLPTVSAGSTICSGIRT
AATVIRKRYPTDSEIIVLTDEEDNTISSCFDLVQSGAIHHTVALGPAARLEBOLS
KMTGLOQYSDQVONGVLDFAALSSGNALIAHISOLESGVNLGNOMMNSVI
YDSVSGDPLFTITHTPTTFEIMDPGSEVNGEILDTTAVILOYGTAKVGMK
YSIOASSOTILITVSRASATLPIITVPPVKNKGRPSPTVYASIRGASPLIR
ASVTLIISVKNKTYTLELDNGAGADATKNGVSRFFATDANGRSVKIHALGCV
TSDRRAAPPKRMYIDGWIEDGEVRMPPEPESYODKOLCFSRSSGSPVATN
VPAAPIPDLFPPCQITDILKASIOGNLVNLWTPAGDYDGRASNYIIRNSTIIVD
LHDHNTSLQVWTTGLIPKEASSEIEFEELGNTFNGCTDIFIAIOAVDKSNLSEI
SNIRAVSEIIPAEOPPEIPEDSPPCPDISINSTIGIHLKIMMWLGLGEMQYTLGIH"

BASE COUNT 860 a 718 c 693 g 666 t

ORIGIN

Query Match 64.8% Score 123.8; DB 10; Length 2937;
Best Local Similarity 78.0%; Pred. No. 2.7e-29;
Matches 149; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 GGCTTTGTAGTGACAAACCAACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAAG 60
DB 1638 GGTATTATCTACTACACCAACCACTAAGTGCGCTTACCTCCAAATCCAGGCAAGCTAAG 1697
QY 61 GTTGGACATGGAATACGATCTGCAAGCAAGCTCAACCAACCTGACCCGACTGTCACG 120
DB 1698 GTTGGCTTTGGAAATACGATCTGCAAGCAAGCTCAACCAACCTGACCCGACTGTCACG 1757
QY 121 TCCCGTGCCTCAATGCTACCTGCTCCATTAAGTACTGCTCAACCAAGCAAGCAAGAC 180
DB 1758 TCCCGTGCAGCAAGTCTACCTGCTCCATTAAGTACTGCTCAACCAAGCAAGCAAGAC 1817
QY 181 ACCAGCAATTT 191
DB 1818 ACAGGGAATTT 1828

RESULT 10
AK000138
LOCUS Homo sapiens cDNA FLJ20131 fis, clone COL06357. PRI 22-FEB-2000
DEFINITION AK000138
ACCESSION AK000138
VERSION AK000138.1 GI:7020030
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_11b:COL clone:COL06357.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (sites)

REFERENCE
AUTHORS Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
Unpublished (2000)

TITLE
JOURNAL
AUTHORS 2 (bases 1 to 1895)
Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission

COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing; Research Association for Biotechnology: cDNA library
construction; 5'-3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
Fax: 81-3-5449-5416
E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286.

FEATURES
Source
1. 1895
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL06357"

/clone_11b="COL"
/issue_type="colon"
/note="cloning vector pME18SFL3"
66..1475
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA90969.1"
/translation="MSKITGSGSHFYVSDPAONNGLLDAFALTSNGTDLISOKSLES
KGLTINSNMAMDVYIIDSTVGKDFEFILITMSLSPSTISMDPSGTIENFTVDAISK
MAYLISPGAKYGTANYNQAANPETILITYTSRAANSVPITVNAKMKDNVSP
SPMIVAEILQGVPELVGANVAFIESONGHREVELELDNGAGADSFNDGVYSRFT
AVTENGERSLKVRAHGANTARLKLRLPLNRAAYIPGVVNGEIEANPREPIDDTQ
TYLEDFSRSGAAGVAVSQPSLPLPDQYPPQIITDLDATVHEDKIILTPAPGNFD
VGKVOYRIIRISASILDSDPDALOVWTTDLSPEAKSKSFAPKPISEENATH
IFAIKSIDKSNLTSKVSINIAQVLEIFIQANDDIDPPTPTPTPDKSKNSGVNSTL
VLSVGSVYIVNFILITSTI"

BASE COUNT 658 a 352 c 357 g 528 t

ORIGIN

Query Match 44.0% Score 84; DB 9; Length 1895;
Best Local Similarity 68.6%; Pred. No. 2e-16;
Matches 133; Conservative 0; Mismatches 55; Indels 6; Gaps 1;

QY 4 TTTGTAGTGACAAACCAACCAAAATGGCTTACCTCCAAATCCAGGCAATGCTTAAGGTT 63
DB 348 TTCACAGTGTATGCACTTCCAAATGGCTTATCTCAGTATTCAGAGAACTGCAAGAGTG 407
QY 34 GGCACATGGAATACAGTCTGC-----AAGCAAGCTCACAAACCTTGACCTGACTGTC 117
DB 408 GGCACATGGAATACAGTCTGC-----AAGCAAGCTCACAAACCTTGACCTGACTGTC 117
QY 118 AGCTCCGCTGCTCAATGCTACCTGCTCCATTAAGTACTGCTCAACCAAGCAAGCAAG 177
DB 498 ACTTCTGACAGCAAAATTTCTGTGCTCCATTAAGTACTGCTCAACCAAGCAAGCAAG 527
QY 178 GACACGACCAATTT 191
DB 528 GACGTAACAGTTT 541

RESULT 11
AF127035
LOCUS Homo sapiens cDNA FLJ20131 fis, clone COL06357. PRI 11-AUG-1999
DEFINITION AF127035
ACCESSION AF127035
VERSION AF127035.1 GI:5726288
KEYWORDS human.
SOURCE Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 3204)

REFERENCE
AUTHORS Agnel, M., Vernal, T. and Culouscou, J.-M.
Identification of three novel members of the calcium-dependent
chloride channel (CaCC) family predominantly expressed in the
digestive tract and trachea
FEBS Lett. 455 (3), 295-301 (1999)

TITLE
JOURNAL
AUTHORS 2 (bases 1 to 3204)
Agnel, M. and Culouscou, J.-M.
Direct Submission

COMMENT
Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des
Carrières, Rueil-Malmaison 92500, France
Location/Qualifiers
1. 3204
/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="colon"

FEATURES
Source
1. 3204
/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="colon"

FEATURES
Source
1. 3204
/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="colon"

CDS

/gene="CaCC2"
29..2782
/gene="CaCC2"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 2"
/db_xref="GI:5726289"
/db_xref="F01:5726289"
/translation="MGLEFRGFVFLVLCILHOSNTSFILNNGEDIVIVIDPSVE
DEKIEQEDWVTFVSTIYFENKREFFKVFYIPENKMEPPYKPKHNEHKAH
VIVAPPLRGDEPTKOTEGEKEGYIHFTPOLLCCKENYEPKPKLYHEHAHL
RMGVDEYNEDPTFRASKIEATRCASGSGNRNRYKCGGSLCSACRISTDTKL
YKDCQFPPDVVOTERKASIMFQSTIDSVETECNEFTNHOEASLQNICNFRSTVEI
SNSBDFKLTIPMTPTPPPPVSLKIRQIVOLYIDKSGSMGKDLRMNOAAHFL
LOTVENSVMGAVHEDSTATVYNKLIQIKSSIDERNTLMAGLPTPLGTSICSGYA
FOVIGELHSDGSEVLTIDGEDNTASSCIDEVQSGAIVFIALGRADVAEIVMS
KITGSHFYVDEQNNGLIDAFGLTSGNDLSKQSLSKGLTINSNAMNDTVI
IDSTYKDTFFLITWNSLPPTISIMDESGTINENTVATISMAVLSIPGAKVGTMA
YNLAKANPELTITVTSRAANSVPIITVAKNMKNVDNSPSPMIVIAELQGVVY
LGANTAFIESQNGHTEVLELDNGAGDSFRKNDGVSRFYATYENGRYSILKVRAG
GANTARLRLRPPLNRAAYIPGVYNGEIEANPREIDEDQTTLEDSRAGGAFV
VSOVSLPLPOYPSOITDADYVHEDKITLWTAPGDNDVGVORVYITRISAIL
DLRSDFDALQVNTDLSKPKANSKESAFRENSFNEMATHIFIAISIDKSNITSK
VSNIAQVTLFIQANPPDIDPTPTPTPPDKSHNSGVNISTLVISGVIVNFTLS
TTI"

BASE COUNT 1098 a 594 c 633 g 879 t
ORIGIN

Query Match 44.0%; Score 84; DB 9; Length 3204;
Best local Similarity 68.6%; Pred. No. 2e-16;
Matches 133; Conservative 0; Mismatches 55; Indels 6; Gaps 1;

OY 4 TTGTGATGACAAAACCAAAATGGCTCTCCAAATCCAGGCGATTGTAAGTT 63
DB 1655 TTCACAGTGAATGCAACTCCAAATGGCTCTCTAGATTCAGGAGCAAGTG 1714
OY 64 GGCACCTGGAATTAAGTCTGC-----AGCAAGCTCAAAACCTTGACCTGCTGC 117
DB 1715 GGCACCTGGGCAATCAATCTCAAGCAAGCAAGCAAGCAAGCAATTAAGTCAAGTA 1774
OY 138 AGCTCCGCGCTCCCAATGCTACCTGCTCCAAATTAAGTCAAGTCAAGCAAG 177
DB 1775 ACTTCTCGAGCAGCAATTTCTTGCTGCTCCAAATCAAGTCAAGTCAAGTAAG 1834
OY 178 GACACCAAGCAATTT 191
DB 1835 GACGTAACAGTTT 1848

RESULT 12
AK000072 3221 bp mRNA PRI 22-FEB-2000
LOCUS AK000072 Homo sapiens CDNA FJ20065 fls. clone COL01613, highly similar to
DEFINITION ECJC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.
ACCESSION AK000072
VERSION AK000072.1 GI:7019922
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon CDNA to mRNA, clone lib.col clone:COL01613.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Kawabata,A., Hiki,j., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDD human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 3221)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio

Sugano, Institute of Medical Science, University of Tokyo, Department
of Virology; Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

COMMENT
NEDD human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'-6' and one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES

source
1..3221
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01613"
/clone_lib="COL"
/issue_type="colon"
/note="Cloning vector pME18SFL3"
1..3221
/note="highly similar to ECJC_BOVIN EPITHELIAL CHLORIDE
CHANNEL PROTEIN"

BASE COUNT 1105 a 600 c 634 g 882 t
ORIGIN

Query Match 44.0%; Score 84; DB 9; Length 3221;
Best local Similarity 68.6%; Pred. No. 2e-16;
Matches 133; Conservative 0; Mismatches 55; Indels 6; Gaps 1;

OY 4 TTGTGATGACAAAACCAAAATGGCTCTCCAAATCCAGGCGATTGTAAGTT 63
DB 1659 TTCACAGTGAATGCAACTCCAAATGGCTCTCTAGATTCAGGAGCAAGTG 1728
OY 64 GGCACCTGGAATTAAGTCTGC-----AGCAAGCTCAAAACCTTGACCTGCTGC 117
DB 1729 GGCACCTGGGCAATCAATCTCAAGCAAGCAAGCAAGCAAGCAATTAAGTCAAGTA 1788
OY 138 AGCTCCGCGCTCCCAATGCTACCTGCTCCAAATTAAGTCAAGTCAAGCAAG 177
DB 1739 ACTTCTCGAGCAGCAATTTCTTGCTGCTCCAAATCAAGTCAAGTCAAGTAAG 1848
OY 178 GACACCAAGCAATTT 191
DB 1839 GACGTAACAGTTT 1862

RESULT 13
AX092338 3265 bp DNA PAT 21-MAR-2001
LOCUS AX092338 Sequence 69 from Patent WO0116318.
ACCESSION AX092338
VERSION AX092338.1 GI:13444483
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
Wood,W.I.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0116318-A 69 08-MAR-2001;
Genentech, Inc. (US)

FEATURES
source
1..3265
Location/Qualifiers

BASE COUNT 1159 a 596 c 632 g 878 t
ORIGIN


```

FEATURES
source
location/Qualifiers
1..2820
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="3"
/tissue_type="lung"
/cell_type="aortic endothelial cells"
194..2656
/note="endothelial adhesion molecule; chloride channel"
/codon_start=1
/product="Lu-ECAM-1"
/protein_id="AAB6531.1"
/db_xref="GI:263767"
/translation="MVLCNLNVILFLTHLLPLPKMKSSVNLINNGDSIVIAINPSVE
DEKLINIKEMVTEASTYLLPHATKRARRYFNNSLLIPMTKKSSEYFIPOESTDQAD
VIANPYKADDEPYLYOAGRCGERKQYIHFTPEFLITNNFHHYSGRVEFVEMHAL
RMGIFEDYNDODPYIYIRKNTIEKTSHTGTHINVEKKCGPGSCITSLCRDSQFG
LYEAKTFPLKKSQTKAKESIMPMPSLASHSNKCYECCETKHNAPLQDKMCKNGKSTG
LIMNSVDQNTSPMTENMPPTHPTEFLSKRSQRYCVLVDKSGSKSADRLPOMNAQD
LYLVIOLEKSLVGMVTFDSVAELIQNH,TRITDQNVYQK,ITAKLPQVANGSTISCRLL
KAGQAOALIHSDQSTSGSEIILLFDGEDNEINSECFEDVKSQALIIHTALPSPAKELE
TLSMNGGGRPEFAKDKITGLTNAPSRISNSGSGITQOALIOLESKALITGRKARGVTV
PVDSTVANDQFFVYVWTWIKQDEIYLODPKKKYATSPKEDKLIKRSARLOIPGIAEF
GWTMYSLLNNHASSQMLVAYVTRBARPTIPPVATAPMHSQHTAHYSPDMIVYQVSG
GHFLPVLGISVAILLETEDGHQVITLMDWNGGRQVYANDGTSIFPDIYGGNGATSLK
VHAQARNRATLNLNROPQNVKPLVYGVGVENKRIILNPREPVKDLAKAKIEDSRLLT
SGGFTYSGADPPPNHSPSVPPSKITDLEAFKEDYIQLSWTAPGVLMDKGKASGSEFF
MSRSHQVAVYLEIOLQHSQSF"
BASE COUNT
940 a 565 c 560 g 755 t
ORIGIN

```

Query/March	18.7%	Score 35.8;	DB 4;	Length 2820;
Best local similarity	57.4%;	Pred. No. 0.77;		
Matches 89;	Conservative 0;	Mismatches 57;	Indels 9;	Gaps 1;
QY	37	CTCCAAATCCAGGCAATGCTAAGGTGGCACTTGGAAATACAGTCT-----GCAA	87	
DB	1835	CTGCAAATACCTGCGATTGTCACAGAGACAGTACTTGGACTTACAGGCTTCTAATAATATCAT	1924	
QY	38	GCAGAGCTACAAACCTTGACCTGACTGTGCACAGTCCCGGCGTCCAAATGCTACCTGGCT	147	
DB	1925	GGCACTCTCAAAATCTATACAGTGCACAGTGCACACTCGAGCAAGAAGTGTCTACTATACCC	1984	
QY	1:8	CCAAATTACAGTACTTCCAAAACGAACAAAGGACAC	182	
DB	1985	CCAGTATATTGCAACAGCTCATGATGATCAACTAC	2019	

Search completed: April 4, 2002, 03:02:50
Job time: 60452 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:35 ; Search time 16681 Seconds
(without alignments)
151.385 Million cell updates/sec

Title: US-09-049-696-12
Perfect score: 235
Sequence: 1 GACACGACGCAAAATTCGCCAG.....CAGTGTAAAGCTGGCGGCTC 235

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gp_est1:*
11: gp_est2:*
12: gp_hic:*
13: gp_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_fod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	233.4	99.3	653	10	AW361534 QV2-CT026
C 2	222.2	94.6	633	10	AW753451 QV2-CT026
C 3	198.2	84.3	496	10	AW753448 QV2-CT026
4	176.4	75.1	450	10	AA734088 vs19c08.r
5	176.4	75.1	479	10	AA541829 vj01f01.r
6	176.4	75.1	911	11	BF578693
7	176.4	75.1	2915	12	AK007466 Mus muscu
8	176.4	75.1	2933	12	AK008659 Mus muscu
9	165.4	70.4	935	11	BF578919
10	162.2	69.0	774	11	BF583551
11	154.4	65.7	974	11	BC962464
12	152.2	64.8	469	10	BE136935

13	140.2	59.7	906	11	BF581342
14	116.8	49.7	703	13	AZ022011
15	116.4	49.5	566	11	BG087141 H3136A06-
16	113.8	48.4	549	10	BE281945 601100284
17	113.2	48.2	470	10	A1197242 u081f01.r
18	113.2	48.2	504	11	BF178212
19	113.2	48.2	844	11	BF174331
20	113.2	48.2	959	11	BF100527
21	109	46.4	455	11	BF083321
22	108.8	46.3	469	10	AA611564
23	107.4	45.7	528	11	BF083270
24	107.4	45.7	647	11	BF678906
25	106.2	45.2	914	11	BG172513
26	104.6	44.5	583	10	AW230825
27	104.4	44.4	917	11	BG172316
28	102.2	43.5	796	11	BF157217
29	98.2	41.8	625	10	AW230824
30	98.2	41.8	1107	11	BF581120
31	93.2	39.7	264	10	BB584478
32	93	39.6	356	10	AW361528
33	83.8	35.7	866	11	BF155803
34	79.2	33.7	286	10	BE715538
35	78.4	33.4	192	10	AW843764
36	77.8	33.1	940	11	BG678187
37	72.6	30.5	770	10	BE738113
38	71.6	30.5	539	13	AZ464851
39	65.8	28.0	420	10	A1641868
40	64.2	27.3	653	10	AW212836
41	60	25.5	717	11	BE850924
42	59.6	25.4	524	10	AA871630
43	58.2	24.8	824	11	BF109423
44	57.2	24.3	490	10	BE217974
45	57	24.3	598	11	BE927204

ALIGNMENTS

RESULT 1
LOCUS AW361534/c
DEFINITION QV2-CT0261-261099-011-f01 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW361534
VERSION AW361534.1 GI:6866288
KEYWORDS EST.
SOURCE human:
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 653)
HCGP <http://www.ludwig.org.br/ORESSES>.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?rl=QV2&rl=QV2-CT0261-261099-011-f01&rl=1999-10-26&rl=1>)
Seq primer: puc 18 forward
High quality sequence start: 37
Location/Qualifiers
1..653
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"


```

/dev_stage="Adult"
/node="Organ: colon; Vector: puc18; Site_1: Smal; Site_2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      140 a      156 c      166 g      191 t
ORIGIN

```

	Query Match	Similarity	99.3%	Score 233.4	DB 10	Length 653
	Best Local	Similarity	99.6%	Pred. No. 1.8e-59		
	Matches 234	Conservative	0	Mismatches	1	Indels 0; Gaps 0
QY	1	GACACCAGCAAAATTCGCCAGAGCCTCTGGTAGTTATGCAAATATTCGCCAGAGCCTCC	60			
DB	340	GACACCAGCAAAATTCGCCAGAGCCTCTGGTAGTTATGCAAATATTCGCCAGAGCCTCC	281			
QY	61	CCAAATTCAGGGCCAGGTGCACAGCCCGATGTAATCGTGAATGGAAAAACAGTTACC	120			
DB	280	CCAAATTCAGGGCCAGGTGCACAGCCCGATGTAATCGTGAATGGAAAAACAGTTACC	221			
QY	121	TTTGAACACTAGTGAATATGAGACAGGTGCTGATGCTACTAAGAGTACGGTGTCTACTCA	180			
DB	220	TTTGAACACTAGTGAATATGAGACAGGTGCTGATGCTACTAAGAGTACGGTGTCTACTCA	161			
QY	181	AGGTAATTCACAACTATATGACACGAATGGTAGTATAGTAAATATGGGGGCTC	235			
DB	160	AGGTAATTCACAACTATATGACACGAATGGTAGTATAGTAAATATGGGGGCTC	106			

RESULT	2
AM753451/c	
LOCUS	
DEFINITION	633 bp mRNA
ACCESSION	OZ-CT0261-26109-011.d05 CT0261 Homo sapiens EST
VERSION	AM753451.1 GI:7666383
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 633)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202653
COMMENT	Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@judwig.org, br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.judwig.org.br/scripts/gethtml2.pl?tbl=OV2&tbl2=OV2-CT0261>
261099-011-d05&ts=1999-10-26&tf=1)
Seq primer: puc 18 forward
High quality sequence start: 66
High quality sequence stop: 90.

```
FEATURES      Location/Qualifiers
source        1..633
              /organism="Homo sapiens"
```

```

/db_xref="taxon:9606"
/clone_lib="Cr0261"
/dev_stage="Adult"
/notes="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 199
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      136 a      151 c      161 g      185 t
ORIGIN

```

	Query Match	Similarity	Score	222.2	DB	10	Length	633
Best Local	Similarity	96.6%	Pred. No.	3.9e-56				
Matches	227	Conservative	0	Mismatches	8	Indels	0	Gaps
QY	1	GACACGACCAATTTCCCGACGCTCTGTAGTTTATGCAAAATATTTGCCACGAGACCTCC	60					
D3	235	GACACGACCAATTTCCCGACGCTCTGTAGTTTATGCAAAATATTTGCCACGAGACCTCC	226					
QY	51	CCATTCTCGGGCCAGTGTCAACGCCCGATTGAATCAGTAATGGAAAAACGTTACC	120					
D3	225	CCATTCTCGGGCCAGTGTCAACGCCCGATTGAATCAGTAATGGAAAAACGTTACC	126					
QY	121	TTGAACTACTGATTAATGAGCAGGCTGTACTTAAGATGACGGTGTCTACTCA	180					
D3	165	TTGAACTACTGATTAATGAGCAGGCTGTACTTAAGATGAGCGGTGTCTACTCT	106					
QY	131	AGGATTTTTCACACTTATGACACGAATGGTAATACAGTGTAAAAAGTGGGCTC	235					
D3	105	ATGATATTCACGAATTATGACACGAATGGTAATACAGTGTAAAAAGGCGGCTC	51					

RESULT	3
AM753444/c	
LOCUS	AM753448 496 bp mRNA
DEFINITION	QV2-CT0261-261099-011-d04 CT0261 Homo sapiens CDNA, mRNA sequence
ACCESSION	AM753448
VERSION	AM753448.1 GI:7668360

REFERENCE	SOURCE	ORGANISM
1 (bases 1 to 496)	human	Homo sapiens
		Chordata; Vertebrata; Euteleostomi;
		Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS: Santos Neto, E., Garcia Collette, R., Veljovski, Almeida, S., Bionhe, M., Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunsteil, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
Sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
COMMENT 20202663
Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, São Paulo-SP,
Brazil
Tel.: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@uclwug.org, br
This sequence was derived from the FAPESP/ICR Human Cancer Genome
project. This entry can be seen in the following URL
(<http://www.uclwug.org.br/scripts/gethtml2.pl?l=VR2-0V2-Cr0261>
261099-011-d04a3=1999-10-26&a=1)
Seq primer: puc 18 forward
High quality sequence start: 59
High quality sequence stop: 89.

FEATURES

LOCUS BF583551 774 bp mRNA EST 12-DEC-2000
DEFINITION 60209604F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4216728 5',
ACCESSION BF583551
VERSION BF583551.1 GI:11657269
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 774)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@email.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM9794 row: g column: 01
High quality sequence stop: 450.
Location/Qualifiers
1. 774
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4216728"
/lab_host="NCI_CGAP_Co24"
/note="Organ: DH10B (T1 phage-resistant)"
/note="Vector: PCMV-SPOrt6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 253 a 208 c 173 g 140 t
ORIGIN

Query Match 69.0%; Score 162.2; DB 11; Length 774;
Best Local Similarity 83.4%; Pred. No. 3,6e-38;
Matches 196; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 2 ACACGAGCAATTCCTCCAGCCCTGTGATGTAATGCAATATCCGACGAGCCCTCC 61
DB 96 ACACGAGCAATTCCTCCAGCCCTGTGATGTAATGCAATATCCGACGAGCCCTCC 155
QY 62 CAATTTTCAGGCGCAGTGTGACAGCCCTGTGATGTAATGCAATATGAAAAACATTA 121
DB 156 CTATTTTCAGGCGCAGGCTGACAGCCTGTGATGTAATGTAATGAAAAACATTA 215
QY 122 TCGAATCTGATGTAATGAGAGCAGGCTGTGATGTAATGAGAGCAGGCTGTAC 181
DB 216 TGAATATTCAGGTAATGAGAGCAGGCTGTGATGTAATGAGAGCAGGCTGTAC 275
QY 182 GGTATTTCACAACT-TATGACAGCAATGATGATACAGTGAATGAAGTGGGGCTC 235
DB 276 GGTTCCTTACAGCTATTTGATGCAATGTAATGATACAGCCTTCAATATATGGCTC 330

RESULT 11
LOCUS BG962464 974 bp mRNA EST 12-JUN-2001
DEFINITION 602827322F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4982105 5',
ACCESSION BG962464
VERSION BG962464.1 GI:14350101
KEYWORDS EST.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 974)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@email.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM10984 row: 1 column: 06
High quality sequence stop: 636.
Location/Qualifiers
1. 974
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4982105"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: DH10B (T1 phage-resistant)"
/note="Vector: PCMV-SPOrt6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 272 a 236 c 257 g 209 t
ORIGIN

Query Match 65.7%; Score 154.4; DB 11; Length 974;
Best Local Similarity 83.9%; Pred. No. 8,2e-36;
Matches 198; Conservative 0; Mismatches 36; Indels 2; Gaps 2;

QY 2 ACACGAGCAATTCCTCCAGCCCTGTGATGTAATGCAATATTCGACGAGCCCTCC 60
DB 216 ACACGAGCAATTCCTCCAGCCCTGTGATGTAATGCAATATTCGACGAGCCCTCC 275
QY 61 CAATTTTCAGGCGCAGTGTGACAGCCCTGTGATGTAATGCAATATGAAAAACATTA 120
DB 276 CCAATTTTCAGGCGCAGGCTGACAGCCTGTGATGTAATGCAATATGAAAAACATTA 335
QY 121 TTGAATCTGATGTAATGAGAGCAGGCTGTGATGTAATGAGAGCAGGCTGTAC 180
DB 336 CTGAATATTCAGGTAATGAGAGCAGGCTGTGATGTAATGAGAGCAGGCTGTAC 395
QY 181 AGTATTTCACAACTTATGACAGCA-TGTGATGTAATGATGTAATGAAGTGGGGCTC 235
DB 336 AGTATTTCACAGCTTTGATGCAATGTAATGATGTAATGTAATGATGAGCTC 451

RESULT 12
LOCUS BE136935 469 bp mRNA EST 21-JUN-2000
DEFINITION u957412.y1 Barstead bowel MRLR9 Mus musculus cDNA clone
IMAGE:1546510 5' similar to TR:088826 O88826 G08-5 PROTEIN.; mRNA
sequence.
ACCESSION BE136935
VERSION BE136935.1 GI:8599435
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 469)
AUTHORS Maria M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person
B., Swaller, T., Gibbons, M., Page, D., Harney, N., Schur, R., Ritter
E., Kohn, S., Shin, T., Jackson, L., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
TITLE The WashU-NCI Mouse EST Project 1999

Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPci-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.bufileo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.bufileo.edu/orderingframe.htm) or from Resea ch Genetics (inf@resgen.com). BAC end page: http://www.tigr.org/cdb/bac/ends/mouse/bac_end_intro.html
 Plate: 301 row: D column: 9
 Seq primer: SP6
 Clones: BAC ends.

FEATURES

source
 Location/Qualifiers
 1. 703
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPci-23-301D9"
 /clone_lib="RPci-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."
 BASE COUNT 171 a 145 c 170 g 216 t 1 others
 ORIGIN

Query Match 49.7%; Score 116.8; DB 13; Length 703;
 Best Local Similarity 83.1%; Pred. No. 1.3e-24;
 Matches 133; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 2 ACACAGCAAAATTCCTCCAGCCCTCTGCTAGTATGCAAAATTCCTCCAGAGCCCTCCC 61
 Db 567 ACACAGGAAATTCCTCCAGCCCTCTGCTAGTATGCAAAATTCCTCCAGAGCCCTCCC 508
 QY 62 CAATTCAGAGGCGCAGTGTACAGCCCTGATGATGATGATGATGATGATGATGATGAT 121
 Db 507 CTATTCAGAGGCGCAGTGTACAGCCCTGATGATGATGATGATGATGATGATGATGAT 448
 QY 122 TGGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 161
 Db 447 TGGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 408

RESULT 15
 BG087141 566 bp mRNA EST 26-JAN-2001
 LOCUS H3136A06-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 DEFINITION H3136A06 5', mRNA sequence.
 ACCESSION BG087141
 VERSION BG087141.1 GI:12569705
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 566)
 Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
 T.S., Carter, M.G. and Ko, M.S.H.
 Verification and initial annotation of NIA mouse 15K cDNA clone set
 Unpublished (2001)
 Other ESTs: H3136A06-3
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit http://lgsun.grc.nia.nih.gov/cDNA/15K.html for details.
 Plate: H3136 row: A column: 06

Seq primer: -21M13 Reverse
 High quality sequence stop: 566
 POLYA=No.

FEATURES

source
 Location/Qualifiers
 1. 566
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="H3136A06"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) -Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A. 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
 BASE COUNT 187 a 131 c 123 g 125 t
 ORIGIN

Query Match 49.5%; Score 116.4; DB 11; Length 566;
 Best Local Similarity 69.1%; Pred. No. 1.7e-24;
 Matches 159; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 13 CACAGCAAAATTCCTCCAGCCCTCTGCTAGTATGCAAAATTCCTCCAGAGCCCTCCC 62
 Db 425 CACAGCCCAATTCCTCCAGCCCTCTGCTAGTATGCAAAATTCCTCCAGAGCCCTCCC 334
 QY 183 AATTCAGAGGCGCAGTGTACAGCCCTGATGATGATGATGATGATGATGATGATGAT 122
 Db 325 TGTTCTGAGAGCCCAATTCCTCCAGCCCTCTGCTAGTATGCAAAATTCCTCCAGAGCCCT 394
 QY 123 GGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 182
 Db 395 GGAAGCTCTGGAGCAATGAGGCGCAGTGTCTGATGATGATGATGATGATGATGATGAT 454
 QY 183 GATATTCACAACTTATGACACGATGATGATGATGATGATGATGATGATGATGAT 232
 Db 455 ATACTTACAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 504

Search completed: April 3, 2002, 20:27:38
 Job time: 40221 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:54:14 ; Search time 521.98 Seconds
(Without alignments)
101.962 Million cell updates/sec

Title: US-09-049-696-12
Perfect score: 235
Sequence: 1 GACACGACCAATTCCTCCAG.....CAGTGTAAAGTGGCGGCTC 235

Scoring table:
IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 segs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_NA:*
2: /cgn2_6/prodata/1/ina/5a_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5b_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6a_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6b_COMB.seq:*
6: /cgn2_6/prodata/1/ina/6c_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	101.4	43.1	576	4	US-09-385-982-23	Sequence 23, Appl
2	90.4	38.5	611	4	US-09-385-982-27	Sequence 27, Appl
3	65	27.7	878	1	US-08-469-667-8	Sequence 8, Appl
4	65	27.7	878	5	PCT-US95-07289-8	Sequence 8, Appl
5	31.4	13.4	531	2	US-08-809-267-9	Sequence 9, Appl
6	31.4	13.4	531	5	PCT-US95-13662A-9	Sequence 9, Appl
7	29.4	12.5	660	4	US-09-328-111-65	Sequence 65, Appl
8	29.4	12.5	3190	4	US-09-286-9598-1	Sequence 1, Appl
9	29.2	12.4	1764	3	US-09-013-881-13	Sequence 13, Appl
10	27.8	11.8	428	4	US-09-004-731-37	Sequence 37, Appl
11	27.8	11.8	428	4	US-09-004-731-39	Sequence 39, Appl
12	27.8	11.8	428	4	US-08-749-699-37	Sequence 37, Appl
13	27.8	11.8	428	4	US-08-749-699-39	Sequence 39, Appl
14	27.8	11.8	717	4	US-09-004-731-43	Sequence 43, Appl
15	27.8	11.8	717	4	US-09-004-731-45	Sequence 45, Appl
16	27.8	11.8	717	4	US-08-749-699-43	Sequence 43, Appl
17	27.8	11.8	717	4	US-08-749-699-45	Sequence 45, Appl
18	27.8	11.8	726	4	US-09-032-215-49	Sequence 49, Appl
19	27.8	11.8	726	4	US-09-032-215-50	Sequence 50, Appl
20	27.8	11.8	841	4	US-09-004-731-40	Sequence 40, Appl
21	27.8	11.8	841	4	US-09-004-731-42	Sequence 42, Appl
22	27.8	11.8	841	4	US-09-032-215-46	Sequence 46, Appl
23	27.8	11.8	841	4	US-09-032-215-48	Sequence 48, Appl
24	27.8	11.8	841	4	US-08-749-699-40	Sequence 40, Appl
25	27.8	11.8	841	4	US-08-749-699-42	Sequence 42, Appl
26	27.8	11.8	1257	4	US-09-554-225-2	Sequence 2, Appl
27	27.8	11.8	1892	4	US-09-554-225-3	Sequence 3, Appl

28	27	11.5	3280	1	US-08-259-000-4	Sequence 4, Appl
29	27	11.5	3280	1	US-08-729-767-6	Sequence 6, Appl
30	26.8	11.4	1080	1	US-07-885-970A-15	Sequence 15, Appl
31	26.8	11.4	1080	1	US-08-298-687A-15	Sequence 15, Appl
32	26.8	11.4	1080	1	US-08-530-797-15	Sequence 15, Appl
33	26.8	11.4	1080	1	US-08-298-829-15	Sequence 15, Appl
34	26.8	11.4	1080	2	US-08-787-535-16	Sequence 16, Appl
35	26.6	11.3	2097	3	US-09-002-5678-2	Sequence 2, Appl
36	26.6	11.3	4253	2	US-08-691-814B-7	Sequence 7, Appl
37	26.4	11.2	794	3	US-08-788-806-1	Sequence 1, Appl
38	26.4	11.2	2719	3	US-08-706-216-1	Sequence 1, Appl
39	26.4	11.2	36519	3	US-08-923-137-2	Sequence 2, Appl
40	26.2	11.1	3666	2	US-08-682-517-13	Sequence 13, Appl
41	26.2	11.1	3666	2	US-08-682-517-14	Sequence 14, Appl
42	26.2	11.1	4197	2	US-08-682-517-7	Sequence 7, Appl
43	26.2	11.1	4197	2	US-08-682-517-8	Sequence 8, Appl
44	26.2	11.1	5910	1	US-08-195-814-1	Sequence 1, Appl
45	26.2	11.1	7125	1	US-07-745-206A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-385-982-23
Sequence 23, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE/REFERENCE: CCDDA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(576)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-23
Query Match 43.1%; Score 101.4; DB 4; Length 576;
Best local similarity 73.7%; Pred. No. 1.3e-26;
Matches 129; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
11 CCAATTTCAGGGCGACATGTCACAGCCCTGATTAACAGATGGAAGAAACAGTTACC 120
2 cctgtctcttgagccatctgacgtctcttcattcaacacagatgacacagaaagt 61
11 TTGGACACTAGTGAATATGAGACAGTGTCTGATCTAAGATGACGATGCTACTCA 180
2 ttggaacttttgataatgtaagtgacggcgatcttcaagaatgtaagtgatctacc 121
11 AGGATTTTCAACAATTATGACACGAATGTAAGTATACAGTGAAGAGCGGCTC 235
2 aggatatttcacagataacagaaatgacagatataagcttaaaagtctgagctc 176
RESULT 2
US-09-385-982-27
Sequence 27, Application US/09385982


```
? Patent No. 6262334.
? GENERAL INFORMATION:
? APPLICANT: ENDEGE, WILSON O., ET AL.
? TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
? TITLE OF INVENTION: PRODUCTS: II
? FILE REFERENCE: CCDNA-260XX
? CURRENT APPLICATION NUMBER: US/09/385,982
? CURRENT FILING DATE: 1999-08-30
? EARLIER APPLICATION NUMBER: 09/328,111
? EARLIER FILING DATE: 1999-06-08
? EARLIER APPLICATION NUMBER: 60/117,393
? EARLIER FILING DATE: 1999-01-27
? EARLIER APPLICATION NUMBER: 60/098,639
? EARLIER FILING DATE: 1998-08-31
? NUMBER OF SEQ ID NOS: 544
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 27
? LENGTH: 611
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(611)
? OTHER INFORMATION: n = A,T,C or G
? GS-09-385-982-27
```

Query Match 38.5%; Score 90.4; DB 4; Length 611;
Best Local Similarity 73.3%; Pred. No. 1.1e-22;
Matches 129; Conservative 0; Mismatches 46; Indels 1; Gaps 1;

OY 61 CCAATTCACAGGCGCAGTGCACAGCCCTGATTGAATCATGTGAAT-GGAAAAACAAGTTAC 119
|| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 ccgtcttcgtgagccaatgtagcttgcattcattgatcacagaatgggacatacatcaagt 61

OY 120 CTTGAACTACTGGTAATATGAGCAGCGTGCATGCTCTAAGAGATGACGGTGTCTAC 179
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 ttggaacttttggaataatgtagcgagcgatcccttaacaatatgatgagttactc 121

OY 180 AAGGATTTTCACACATTATGACACGAATGGTAGATACTGTA AAAAGTGGGGCTC 235
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 cagatattttaagatatatacagaanaatgcsagataatgcttaaagaattcgggctc 177

RESULT 3
US-08-469-667-8
Sequence 8, Application US/08469667
Patent No. 5733748
GENERAL INFORMATION:
APPLICANT: YU, Guo-Liang
APPLICANT: Rosen, Craig
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gillfillan, Cecchi,
ADDRESSER: Stewart & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469, 667
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-435
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 878 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..685
DS-08-459-667-8

```

Query Match      27.7%; Score 65; DB 1; Length 878;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

```

07 TGTCTACTCAAGGATTTTCACAACTTATGCACGAAGTAGTGATACGGTAAAGCGC 230
09 |||||
D9 1 TGTCTACTCAAGGATTTTCACAACTTATGCACGAAGTAGTGATACGGTAAAGCGC 60
-231 GCCTC 235
QY |||||
D9 51 GGCTC 65

RESULT : 4
 PCT-US95-07289-8
 Sequence 8, Application PC/TUS9507289
 GENERAL INFORMATION:
 APPLICANT: Yu, Guo-Liang
 APPLICANT: Rosen, Craig
 TITLE OF INVENTION: Colon Specific Genes and Proteins
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
 ADDRESSEE: Stewart & Olstein
 STREET: 6 Becker Farm Road
 CITY: Roseland
 STATE: NJ
 COUNTRY: USA
 ZIP: 07068-1739
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT-US95/07289
 FILING DATE: 06-JUN-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Ferraro, Gregory D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 325800-265
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 878 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..605
 PCT-US95-07289-8

Query Match 27.7%; Score 65; DB 5; Length 878;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 171 TGTCTACTCAAGGTTATTCACCACTATGACAGCAATGGTAGATACAGTCTAAAGTGGC 230
|||||
DB 1 TGTCTACTCAAGGTTATTCACCACTATGACAGCAATGGTAGATACAGTCTAAAGTGGC 60
QY 231 GGCTC 235
|||||
DB 61 GGCTC 65

RESULT 5
US-08-809-267-9
Sequence 9, Application US/08809267
Patent No. 5861296
GENERAL INFORMATION:
APPLICANT: LENNOX, Tricia L.
APPLICANT: SLATKO, Barton E.
APPLICANT: SEARS, Lauren E.
TITLE OF INVENTION: PURIFIED THERMOSTABLE INORGANIC
TITLE OF INVENTION: PYROPHOSPHATASE OBTAINABLE FROM THERMOCOCCUS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEW ENGLAND BIOLABS, INC.
STREET: 32 TOZER ROAD
CITY: BEVERLY
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01915
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,267
FILING DATE: 12-MAR-1997
CLASSIFICATION: 433
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13662
FILING DATE:
APPLICATION NUMBER: US 08/329,721
FILING DATE: 25-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, GREGORY D.
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-105-PCT
TELEPHONE: (508) 927-5054
TELEFAX: (508) 927-1705
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 531 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-809-267-9

Query Match 13.4%; Score 31.4; DB 2; Length 531;
Best Local Similarity 56.2%; Pred. No. 0.086;
Matches 59; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 8 GCAAAATCCCGCCCTCTGATGATATGCAATATTCGCCAAGAGGCTCCCAATTC 67
|||||
DB 262 GCAAGACCAATAGGCTCTTCAAGATGATAGACAGCGCGCAAGGACTCAAGATTTG 321

QY 171 TGTCTACTCAAGGTTATTCACCACTATGACAGCAATGGTAGATACAGTCTAAAGTGGC 112
|||||
DB 1 TGTCTACTCAAGGTTATTCACCACTATGACAGCAATGGTAGATACAGTCTAAAGTGGC 366

RESULT 6
PCT-US95-13662A-9
Sequence 9, Application PC/TUS9513662A
GENERAL INFORMATION:
APPLICANT: LENNOX, Tricia L.
APPLICANT: SLATKO, Barton E.
APPLICANT: SEARS, Lauren E.
TITLE OF INVENTION: PURIFIED THERMOSTABLE INORGANIC
TITLE OF INVENTION: PYROPHOSPHATASE OBTAINABLE FROM THERMOCOCCUS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEW ENGLAND BIOLABS, INC.
STREET: 32 TOZER ROAD
CITY: BEVERLY
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01915
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13662A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,721
FILING DATE: 25-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, GREGORY D.
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-105-PCT
TELEPHONE: (508) 927-5054
TELEFAX: (508) 927-1705
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 531 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
PCT-US95-13662A-9

Query Match 13.4%; Score 31.4; DB 5; Length 531;
Best Local Similarity 56.2%; Pred. No. 0.086;
Matches 59; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 8 GCAAAATCCCGCCCTCTGATGATATGCAATATTCGCCAAGAGGCTCCCAATTC 67
|||||
DB 262 GCAAGACCAATAGGCTCTTCAAGATGATAGACAGCGCGCAAGGACTCAAGATTTG 321

Query Match 13.4%; Score 31.4; DB 2; Length 531;
Best Local Similarity 56.2%; Pred. No. 0.086;
Matches 59; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 8 GCAAAATCCCGCCCTCTGATGATATGCAATATTCGCCAAGAGGCTCCCAATTC 67
|||||
DB 262 GCAAGACCAATAGGCTCTTCAAGATGATAGACAGCGCGCAAGGACTCAAGATTTG 321

Query Match	Similarity	Score	DB	Length
Best Local	52.9%	Pred. No. 1.1		3190
Matches	63	Conservative	0	Mismatches 56; Indels 0; Gaps 0.
0y	82	ACAGCCCGTATTGAATCACTGATGTAATGAGAAAACAGTACTTGGAGTAATCGATATGATATGA	141	
db	1215	aaccaccataaagaacaccagggatgtaagaataaagaatacccttttgaacctgctctggg	1274	

693 AATGAAAACAGTTACTCTTGAACTACTGATTAATGACGAGGTGCTGATGCTACTAAG 162
 693 AATGAGTAGCAAAACCTCTGGAAACCAATGCGTAGATGAAGAAGCTAATTAAGTACTTTAAG 6322
 693 GATGACGTGTCTACTCAAGTATTTTCACACTATATGACAGATGCTAGATACAGTGA 222
 693 AAGGTGTTTTTATTACAGAAAGTTACAACTATATCCAACTTAGATCATCATGCTA 692
 693 AA 224
 693 AA 694

RESULT 10
US-09-004-731-37
Sequence 37, Application US/09004731
Patent No. 6177258
GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,731
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 428 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..428
FEATURE:
NAME/KEY: Xaa = any amino acid
LOCATION: 17
US-09-004-731-37

Query Match 11.8%; Score 27.8; DB 4; Length 428;
Best local Similarity 55.8%; Pred. No. 1.5;
Matches 53; Conservative 0; Mismatches 42; Indels 0; gaps 0;

QY 98 CAGTGAATGGAAGAAAGACTTGTGATGATGATGAGAGAGTGTGATGCTA 157
DB 203 CAGAACTGGAAAGAAAGAGTTCAGAAAGCCGCTGTTAGTGTACAGATGGGGGCTA 268

QY 158 CTAAGATGACGGTGTCTACTCAAGGTATTTCACA 192
DB 269 CTATGTGGGGCCGACGATCAACGATCTATAA 303

RESULT 11
US-09-004-731-39/C
Sequence 39, Application US/09004731
Patent No. 6177258
GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley

APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,731
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 428 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-004-731-39

Query Match 11.8%; Score 27.8; DB 4; Length 428;
Best local Similarity 55.8%; Pred. No. 1.5;
Matches 53; Conservative 0; Mismatches 42; Indels 0; gaps 0;

QY 98 CAGTGAATGGAAGAAAGACTTGTGATGATGATGAGAGAGTGTGATGCTA 157
DB 220 CAGAACTGGAAAGAAAGAGTTCAGAAAGCCGCTGTTAGTGTACAGATGGGGGCTA 161

QY 158 CTAAGATGACGGTGTCTACTCAAGGTATTTCACA 192
DB 160 CTATGTGGGGCCGACGATCAACGATCTATAA 126

RESULT 112
US-08-749-699-37
Sequence 37, Application US/08749699
Patent No. 6210920
GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 428 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..428
FEATURE:
NAME/KEY: Xaa = any amino acid
LOCATION: 17
US-08-749-699-37

Query Match
Best Local Similarity 55.8%; Score 27.8; DB 4; Length 428;
Matches 53; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 98 CAGTGAATGAAAAACAGTCTGACTGATGATGAGAGAGTGTGATGCTA 157
DB 209 CAGAACTGAAAAAGAGTTCAGAAAGCGGCGACTGCTAGTGTGACAGAGATGGGGGCTA 268
QY 158 CTAAGATGAGCGGTCTACTCAAGGATTTTACA 192
DB 269 CTATGTGGGGGCCAGCATCAACGATCTAATAA 303

RESULT 13
US-08-749-699-39/C
Sequence 39, Application US/08749699
Patent No. 6210920
GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 428 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-749-699-39

Query Match
Best Local Similarity 55.8%; Score 27.8; DB 4; Length 428;
Matches 53; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 98 CAGTGAATGAAAAACAGTCTGACTGATGATGAGAGAGTGTGATGCTA 157
DB 220 CAGAACTGAAAAAGAGTTCAGAAAGCGGCGACTGCTAGTGTGACAGAGATGGGGGCTA 161
QY 158 CTAAGATGAGCGGTCTACTCAAGGATTTTACA 192
DB 150 CTATGTGGGGGCCAGCATCAACGATCTAATAA 126

RESULT 114
US-09-004-731-43
Sequence 43, Application US/09004731
Patent No. 6177258
GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 717 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

This Page Blank (uspto)

PD 10-SEP-1999. .
XX 03-MAR-1999; 99WO-US04703.
PF
XX 03-MAR-1998; 98US-0076815.
PR
XX
PA (MAGA-) MAGALIN PHARM INC.
XX
XX Holroyd KJ, Levitt RC, Maloy WL, Louhede J, McLane M;
PI Nicolaides NC, Zhou Y, Dong Q;
PI
XX WPL; 1999-550979/46.
DR
DR P-PSDB: AAB74824.
XX
XX New nucleic acid encoding calcium activated chloride channel, used to
PF identify, e.g. specific modulators for treating atopic allergy -
PF
PS Claim 1; Fig 4B; 75pp; English.

CC The present sequence encodes the human interleukin 9 (IL-9) induced
CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
CC have anti-allergic, anti-asthmatic, anti-inflammatory and
CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
CC used to alleviate asthma (or more generally atopic allergy), while those
CC (E) that upregulate ICACC are used to treat cystic fibrosis (CF) and
CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma
CC (increased levels) or IBD (reduced levels), also for monitoring
CC treatment of these conditions. The ICACC proteins can be used:
CC (i) to raise specific antibodies (Ab), useful: (a) as immunoassay
CC reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
CC (A) to disrupt bonding between ICACC and its ligands; and (iii) to
CC identify modulators and binding partners. ICACC polynucleotides can be
CC used to generate transgenic animals or recombinant cells, used to screen
CC for antagonists, also as a source of therapeutic antisense agents or
CC diagnostic probes (for quantifying mRNA expression, e.g. for
CC identification of modulators).
XX
XX Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;

CC recipients; and (b) as therapeutic (AT), (11) as transgene (TG), and (12) as
CC (A) to disrupt bonding between ICACC and its ligands; and (11) to
CC identify modulators and binding partners. ICACC polynucleotides can be
CC used to generate transgenic animals or recombinant cells, used to screen
CC for antagonists, also as a source of therapeutic antisense agents or
CC diagnostic probes (for quantifying mRNA expression, e.g. for
CC identification of modulators).

Query Match	100.0%	Score 235	DB 20	Length 2745
Best local Similarity	100.0%	Pred. No. 7.7e-70		
Matches 235	0	Mismatches 0	Indels 0	Gaps 0

1 GAACACAGCAAAATTCGCCAGCCCTCTGCTAGTTATGCAAAATATTCGCCAAGAGCCTCC 60

Query Match	100.0%	Score 235	DB 20	Length 2745
Best local Similarity	100.0%	Pred. No. 7.7e-70		
Matches 235	0	Mismatches 0	Indels 0	Gaps 0

1 GAACACAGCAAAATTCGCCAGCCCTCTGCTAGTTATGCAAAATATTCGCCAAGAGCCTCC 60

Query Match	100.0%	Score 235	DB 20	Length 2745
Best local Similarity	100.0%	Pred. No. 7.7e-70		
Matches 235	0	Mismatches 0	Indels 0	Gaps 0

1 GAACACAGCAAAATTCGCCAGCCCTCTGCTAGTTATGCAAAATATTCGCCAAGAGCCTCC 60

Query Match	100.0%	Score 235	DB 20	Length 2745
Best local Similarity	100.0%	Pred. No. 7.7e-70		
Matches 235	0	Mismatches 0	Indels 0	Gaps 0

1 GAACACAGCAAAATTCGCCAGCCCTCTGCTAGTTATGCAAAATATTCGCCAAGAGCCTCC 60

Query Match	100.0%	Score 235	DB 20	Length 2745
Best local Similarity	100.0%	Pred. No. 7.7e-70		
Matches 235	0	Mismatches 0	Indels 0	Gaps 0

1 GAACACAGCAAAATTCGCCAGCCCTCTGCTAGTTATGCAAAATATTCGCCAAGAGCCTCC 60

Query Match	100.0%	Score 235	DB 20	Length 2745
Best local Similarity	100.0%	Pred. No. 7.7e-70		
Matches 235	0	Mismatches 0	Indels 0	Gaps 0

1 GAACACAGCAAAATTCGCCAGCCCTCTGCTAGTTATGCAAAATATTCGCCAAGAGCCTCC 60

Query Match	100.0%	Score 235	DB 20	Length 2745
Best local Similarity	100.0%	Pred. No. 7.7e-70		
Matches 235	0	Mismatches 0	Indels 0	Gaps 0

1 GAACACAGCAAAATTCGCCAGCCCTCTGCTAGTTATGCAAAATATTCGCCAAGAGCCTCC 60

expression inhibition; antisense therapy; gene therapy;
 chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
 ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 23..2767
 /tag= a
 /product= "Human CLCA1"
 /transl_except= (pos:476..478, aa:lys)

WO200138530-A1.

31-MAY-2001.

22-NOV-2000; 2000WO-JP08232.

24-NOV-1999; 99JP-0333479.

27-APR-2000; 2000JP-0127589.

(TAKE) TAKEDA CHEM IND LTD.

Nakanishi A, Morita S;
 WPI: 2001-355935/37.
 P-PSDB: AAB73716.

New antisense nucleotide, useful for treatment and prevention of
 bronchial asthma and chronic obstructive pulmonary disease -

Example 5; Page 92-94; 104pp; Japanese.

The invention relates to an antisense nucleotide targeted to the mouse
 Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 the CLCA1 gene (coding sequence shown in AAH46102). The invention also
 relates to an antibody specific for the Gob-5 protein, medical and
 diagnostic compositions containing the antisense nucleotide or the
 antibody, and methods and kits for screening for compounds which inhibit
 the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 The antisense oligonucleotides and antibody are therefore useful for the
 treatment and prevention of bronchial asthma and chronic obstructive
 pulmonary disease. The present sequence represents human CLCA1 cDNA.

Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;

Query Match 100.0%; Score 235; DB 22; Length 2825;
 Best Local Similarity 100.0%; Pred. No. 7.8e-70;
 Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACACCAGCAAAATTCCTCCGCTCTGTGATGTAATTCGCAAGAGGCTCC 60
 1820 gacaccagcaaatTCCTCCGCTCTGTGATGTAATTCGCAAGAGGCTCC 1879

61 CCAATTTTCAGGCGCAGTGTCAAGCCCTGATGTAATGCAATGCAAGAGTACC 120
 1880 ccaatttTCAGGCGCAGTGTCAAGCCCTGATGTAATGCAATGCAAGAGTACC 1939

121 TTGGAACTACTGTGATATGAGCAGGTGCTGATGTAAGAGTACGGTGTCTACTCA 180
 1940 ttggaactactGTGATATGAGCAGGTGCTGATGTAAGAGTACGGTGTCTACTCA 1999

181 AGGTATTTACACACTTATGACAGAGTGTAGATACAGTAAATGGGGGCTC 235
 2000 aggtattTTACACACTTATGACAGAGTGTAGATACAGTAAATGGGGGCTC 2054

RESULT 4
 AAH34879
 ID AAH34879 standard; cDNA; 2854 BP.
 XX
 AC AAH34879;

Human colon cancer antigen encoding cDNA SEQ ID NO:1961.

Human: colon cancer; colon cancer antigen; diagnosis; detection;
 colorectal carcinoma; chromosome 1; ss.

Homo sapiens.

WO200122920-A2.

05-APR-2001.

28-SEP-2000; 2000WO-US26524.

29-SEP-1999; 99US-0157137.

03-NOV-1999; 99US-0163280.

(HUMAN) HUMAN GENOME SCI INC.

Rosen SM, Barash SC, Birse CE, Rosen CA;
 WPI: 2001-235357/24.
 P-PSDB: AAG75474.

Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 useful for preventing, diagnosing and/or treating colorectal cancers -

Claim 1; Page 3462-3463; 9803pp; English.

AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 cancer-associated nucleic acid molecules (N) and proteins (P), where
 the proteins are collectively known as colon cancer antigens. The colon
 cancer antigens have cytostatic activity and can be used in gene
 therapy and vaccine production. N and P may be used in the prevention,
 diagnosis and treatment of diseases associated with inappropriate P
 expression. For example, N and P may be used to treat disorders
 associated with decreased expression by rectifying mutations or deletions
 in a patient's genome that affect the activity of P by expressing P.
 Additionally, N may be used to produce the colon cancer-associated P,
 by inserting the nucleic acids into a host cell and culturing the cell
 to express the proteins. N and P can be used in the prevention, diagnosis
 and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 and AAG77789 represent sequences used in the exemplification of the
 present invention.

N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 missing at time of publication, meaning no sequences are present for
 SEQ ID NO:1027 to 1052, 7921 and 7922.

Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match 100.0%; Score 235; DB 22; Length 2854;
 Best Local Similarity 100.0%; Pred. No. 7.8e-70;
 Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACACCAGCAAAATTCCTCCGCTCTGTGATGTAATTCGCAAGAGGCTCC 60
 1820 gacaccagcaaatTCCTCCGCTCTGTGATGTAATTCGCAAGAGGCTCC 1891

61 CCAATTTTCAGGCGCAGTGTCAAGCCCTGATGTAATGCAATGCAAGAGTACC 120
 1880 ccaatttTCAGGCGCAGTGTCAAGCCCTGATGTAATGCAATGCAAGAGTACC 1951

121 TTGGAACTACTGTGATATGAGCAGGTGCTGATGTAAGAGTACGGTGTCTACTCA 180
 1932 ttggaactactGTGATATGAGCAGGTGCTGATGTAAGAGTACGGTGTCTACTCA 2011

181 AGGTATTTACACACTTATGACAGAGTGTAGATACAGTAAATGGGGGCTC 235
 2012 aggtattTTACACACTTATGACAGAGTGTAGATACAGTAAATGGGGGCTC 2066

XX	AAAF81787	standard; cDNA; 2854 BP.
XX	AAAF81787;	
XX	12-JUN-2001	(first entry)
DE	Human secreted protein gene 1 SEQ ID NO:11.	
XX	Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;	
KW	dermatological; immunosuppressive; antinflammatory; anti-HIV;	
KW	immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;	
KW	ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;	
KW	antialzheimer; antiparkinsonian; antimicrobial; vulnerrary; gene therapy;	
KW	immune disorder; hyperproliferative disorder; cardiovascular disease;	
KW	cancer; angiogenic disorder; neurological disorder; infectious disease;	
KW	wound healing; regeneration; chemotaxis; chromosome 1; ss.	
OS	Homo sapiens.	
XX	MO200112775-A2.	
PN		
XX	22-FEB-2001.	
PD		
PF	16-AUG-2000; 2000MO-US22325.	
XX		
PR	17-AUG-1999; 99US-0149182.	
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;	
PI	Bitse CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;	
XX		
DR	WPI: 2001-147550/15.	
XX	P-PSDB; AAB74733.	
XX		
PT	Nucleic acids encoding 25 human secreted polypeptides, useful for	
PT	preventing, diagnosing and/or treating e.g. cancers, Parkinson's	
XX	disease and diabetic retinopathy -	
PS	Claim 1; Page 441; 485pp; English.	
XX		
CC	AAAF1787 to AAF81817 encode the human secreted proteins given in AAB74733	
CC	to AAB74772. Human secreted proteins can have activities based on the	
CC	ligands and cells they are expressed in. Example of activities include:	
CC	immunomodulatory; antisclerotic; dermatological; immunosuppressive;	
CC	antinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;	
CC	vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;	
CC	anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and	
CC	vulnerrary. Human secreted proteins can be used in gene therapy and	
CC	vaccine. Human secreted protein nucleotide sequences (NAM1) and proteins	
CC	(PEP1) may be used in the prevention, diagnosis and treatment of diseases	
CC	associated with inappropriate polypeptide expression. For example, NAM1	
CC	and PEP1 may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patients genome	
CC	that affect the activity of proteins by expressing inactive proteins or	
CC	to supplement the patients own production of polypeptides. Disorders that	
CC	may be prevented, diagnosed and/or treated include immune disorders,	
CC	hyperproliferative disorders (e.g. cancers), cardiovascular diseases,	
CC	angiotonic disorders, neurological disorders, infectious diseases and/or	
CC	for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to	
CC	AAAF1786 and AAB74732 represent sequences used in the exemplification of	
CC	the present invention.	
XX		
SO	Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;	

Query Match	100.0%	Score 235;	DB 22;	Length 2854;
Best Local Similarity	100.0%	Pred. No. 7.8e-70;		
Matches 235; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

0Y 11 GACACGACGAAATTCCTCCAGCCCTCGAGTATGTTATGCAAAATATTCGCGAAGGAGCCGCC 60
 1691 gacacgacgaaatctccagccctcggatgtatgtatgcaaaatattcgcgaaggagccccc 1891
 1692 gacacgacgaaatctccagccctcggatgtatgtatgcaaaatattcgcgaaggagccccc 1891
 0Y 61 CCAATTCGACGGGCGCAGTGTACACAGCCCTGATTGATCAGTGAATGAAAAACAGTTACC 120
 1892 ccaattctcagggccagtgacacagccctgatgtatcagtgaaatgaaaaaacagttacc 1951
 0Y 121 TTGGAATCTACTGATTAATGAGACAGTGTGCTGATGCTACTACAGACAGCGTCTACTCA 180
 1952 ttggaactactgataataatgagcaggtgtctgatgtactaaagatgacggtgtactca 2011
 0Y 181 AGGATTTTCAACACTTATGACAGCAATGGTATGATACAGTGAAGTCAAGTGGGAGCTC 235
 2012 aggatcttcacacactatgacacgcaatggtatgatacagtgctaaagtcgagctc 2066
 RESULT 6
 AAH33285
 ID AAH33285 standard; cDNA; 2867 BP.
 AC AAH33285:
 XX 03:SEP-2001 (first entry)
 DT
 XX Human colon cancer antigen encoding cDNA SEQ ID NO:341.
 DE Human colon cancer antigen; colon cancer antigen; diagnosis; detection;
 KM colorectal carcinoma; chromosome 1; ss.
 XX Homo sapiens.
 OS
 XX NC020122920-A2.
 PN
 XX 05-APR-2001.
 ED
 XX 28-SEP-2000; 2000MO-US26524.
 XX
 XX 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 XX (HIMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruden SM, Barash SC, Birse CE, Rosen CA;
 FI WFL: 2001-235357/24.
 XX
 DR P-SSDB; AAG73854.
 XX
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 FT useful for preventing, diagnosing and/or treating colorectal cancers -
 FT
 XX
 XX Claim 1; Page 2452-2453; 9803pp; English.
 XX
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inductive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.3. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
CC

XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;
SQ

Query Match 100.0%; Score 235; DB 22; Length 2867;
Best Local Similarity 100.0%; Pred. No. 7.9e-70;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCAGCAAAATTTCCCGACCCCTGTGTAGTTATGCAAAATTTTCGCCAAGGAGGCTCC 60
DB 1836 gacaccagcaaaatccccccgcccctcgtgagttatgcaaaatcccccaagagagccctcc 1895
QY 61 CCAATTTCTCAGGCGCAGTGTCAACGCCCTGATTGAATCAGTGAATGGAATAACAGTTACC 120
DB 1896 ccaattctcagggccaggtgcaacgccctgattgaaatcagtgaaataaacagttacc 1955
QY 121 TTGGAACTACTGTGATATGAGACAGGTGCTGATGCTACTTAAGATGACGGTGTCTACTCA 180
DB 1956 ttggaactactgtgataatgagcaggtgctgacttaagatgacggtgtctactca 2015
QY 181 AGGTATTTCAACAATTATGACAGCAAGTGTAGATACAGTAAAGTGGGGGCTC 235
DB 2016 aggtatttcaacaattatgacacgaatgtagatacagtgtaaagtgcgggctc 2070

RESULT 7
AAH35019
ID AAH35019 standard; cDNA; 3109 BP.

AC AAH35019;
XX
DT 03-SEP-2001 (first entry)

XX Human colon cancer antigen encoding cDNA SEQ ID NO:2101.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma; ss.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HDMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX P-PSDB; AAG73614.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX Claim 1; Page 3587-3588; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing of P,
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell

CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match 100.0%; Score 235; DB 22; Length 3109;
Best Local Similarity 100.0%; Pred. No. 8.1e-70;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCAGCAAAATTTCCCGACCCCTGTGTAGTTATGCAAAATTTTCGCCAAGGAGGCTCC 60
DB 1836 gacaccagcaaaatccccccgcccctcgtgagttatgcaaaatcccccaagagagccctcc 1895
QY 61 CCAATTTCTCAGGCGCAGTGTCAACGCCCTGATTGAATCAGTGAATGGAATAACAGTTACC 120
DB 1896 ccaattctcagggccaggtgcaacgccctgattgaaatcagtgaaataaacagttacc 1955
QY 121 TTGGAACTACTGTGATATGAGACAGGTGCTGATGCTACTTAAGATGACGGTGTCTACTCA 180
DB 1956 ttggaactactgtgataatgagcaggtgctgacttaagatgacggtgtctactca 2015
QY 181 AGGTATTTCAACAATTATGACAGCAAGTGTAGATACAGTAAAGTGGGGGCTC 235
DB 2016 aggtatttcaacaattatgacacgaatgtagatacagtgtaaagtgcgggctc 2070

RESULT 8
AA209840
ID AA209840 standard; cDNA; 3111 BP.

XX AA209840;

XX 26-NOV-1999 (first entry)

XX Human membrane spanning protein MSP-5 cDNA fragment 2.

XX Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;
KM neoplastic disorder; immunological disorder; reproductive disorder;

XX MSP-5; ds.

XX Homo sapiens;

XX WO9946380-A2;

XX 16-SEP-1999.

XX 09-MAR-1999; 99WO-US05073.

XX 13-MAR-1998; 98US-0039064.

XX (INCY-) INCYTE PHARM INC.

XX Tagg YT, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;

XX Guigler KT, Kaser MR, Baughn MR, Shah P;

XX WPI; 1999-551409/46.

XX P-PSDB; AAV33298.

XX New human membrane spanning proteins used to, e.g. prevent and treat
XX neoplastic disorders -
XX Example 1; Page 80-81; 81pp; English.

XX This invention describes novel human membrane spanning proteins (MSPs),
CC and the polynucleotides encoding them. The products of the invention are
CC used to diagnose, prevent and treat neoplastic, immunological and
CC reproductive disorders. This sequence encodes a human membrane spanning

CC protein MSP-5 fragment.
XX
SQ Sequence 3111 BP, 968 A, 664 C, 691 G, 788 T; 0 other

Query Match	100.0%	Score 235	DB 20	Length 311
Best Local Similarity	100.0%	Pred. No.	8	1e-70
Matches 235	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Accession	Sequence	Position
Oy	1 GACACGACGCAAAATTCCTCCAGCCCTCTGCTGTAGTTATGACAAATTCCTGCCAAGACCTCC	60
Db	1831 gacacgacgaatctcccgagccctctgtagtctttagcaaatatctcgcaagagctcc	1890
Oy	61 CCAATTCTCAGGCGCACTGTGCACAGCCCTGATTTGAATCATGTAATGGAATAACACTTACC	120
Db	1891 ccaattctcagggccagtgccacagccctgattgtaactagtgaaatggaataaacagttacc	1950
Oy	121 TTGACACACGCAATGATGAGCAGGCGCATGCACTACAAAGATGACGCTGTCTACACA	180
Db	1951 ttggaactacagataatggaagcagcagtgctgacgacctaagaatgaaagtgctcaacca	2010
Oy	181 AGGATTTTCACAACTATTGACACAGCAATGTGATATACATGTTAAAGTGGCGGCTC	235
Db	2011 aggatcttcaaacactatgacacagcaatgtgagttacagtgataaagtggggctc	2065

RESULT	9
AAI29502	
ID	AAI29502 standard; cDNA; 3311 BP.

AC	AA129502;
XX	
DT	12-OCT-2001 (first entry)
XX	
DE	C902P determined cDNA sequence

KM Human, immunotherapy; diagnosis; colon cancer; colon tumour;
KW immunogenic; gene therapy; vaccine; colonic cancer; ss.

OS	Homo sapiens.
XX	
PN	W0200149716-A2.

PD	12-JUL-2001.
XX	
PF	29-DEC-2000; 2000WO-US35596.

PR	30-DEC-1999;	99US-0476296
PR	10-FEB-2000;	2000US-0480321
PR	15-EBN-2000;	2000US-0504629
PR	06-MAR-2000;	2000US-0519444
PR	19-MAY-2000;	2000US-0575251
PR	29-JUN-2000;	2000US-0609448
PR	28-AUG-2000;	2000US-0649811

(CORIXA) CORIXA CORP.

PI Xu J, Lodes MJ, Secríst H, Benson DR, Meagher MJ, Stolk JA.
PI King GE, Wang T, Jiang Y;
XX
DR WPI: 2001-441847/47.

PT Colon tumor associated proteins and nucleic acids useful for the prevention, diagnosis and treatment of colonic cancer -

PS Claim 2; Page 425-426; 472pp; English

CC The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytosolic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be

used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of TCAPs by expressing inactive proteins or to supplement the patients own production of them. Additionally, (II) may be used to produce the TCAP proteins, by inserting the nucleic acids into a host cell, culturing the cell to express the protein. (II) and its complementary sequences may also be used as DNA probes in diagnostic polymerase chain reaction (PCR) and hybridisation assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. (I) may also be used as antigens in the production of antibodies against TCAPs and in assays to identify modulators of TCAP expression and activity. Anti-(I) antibodies and antagonists may also be used to down regulate TCAP expression and activity. The anti-(I) antibodies may also be used as diagnostic agents for detecting the presence of TCAPs in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). AA182460 to AA129512 and AA124594 to AA124523 represent nucleotide and amino acid sequences given in the exemplification of the present invention.

Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;

Query Match	100.0%	Score 235;	DB 22;	Length 3311;
Best Local Similarity	100.0%	Pred. No. 8,4e-70;		
Matches 235;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GACCGAGCAATTCGCCAGCCCTCGTGTTTGAACAAATTTGCCAAGAGCTCC	60
Ed	219	gacaccagcaaatlccccagccctctgtagtlltaagcaaatlctgcgaagagcctcc	2206
QY	61	CCAAATTCACAGGCGCAGCTGTGCACAGCCCTGATTGAATCACTGATGAGAAACACTTACC	120
Ed	229	ccaattctcagggccagtgltcaacagcccttgatgaatcagtagatggaataaacagtlacc	2266
QY	131	TTGAGCACTACTGGTAATGAGACAGGTGCTGATGCTACTCTAAGAGTACGGGTCTACTCA	180
Ed	229	ttggaactactcgtgataatgagcagcgtgcctgactactaagatgaacggtctctactca	2326
QY	181	AGGATTTCACAACTTATGACAGCAATGATGATACAGTGTAAATGCGGGCTC	235
Ed	239	aggatatttcacaacttatgacaagaatgtagatatacagtgataaagtgcgggctc	2383

RESULT	.10
AAH46121	
ID	AAH46121 standard; cDNA; 486 BP

AC AA46121;
XX
11-SEP-2001 (first entry)

Mouse Gob-5 cDNA hybridisation probe, SEQ ID NO:23.

KM mouse C50-5; murine; goblet cell; human C1CA1 orthologue; drug screening
 KM chondrogenesis inhibition; antisense therapy; gene therapy; bronchial asthma
 KM chronic obstructive pulmonary disease; antiasthmatic;
 KM hybridisation probe; ss.

OS	Misc. sp.
XX	
PN	WO200138530-A1

PD 31 MAY-2001.
XX

XX	24 NOV-1999;	99JP-0333479.
PR	27-APR-2000;	2000JP-0127589.
PR		

(TJKE) TAKEDA CHEM IND LTD.

PI Naikanishi A, Morita S;
XX

DR	WPI; 2001-355935/37.
XX	
PT	New antisense nucleotide, useful for treatment and prevention of
XX	bronchial asthma and chronic obstructive pulmonary disease
PS	
XX	Example 2; Page 91; 104pp; Japanese.
CC	The invention relates to an antisense nucleotide targetted to the mouse
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC	the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC	relates to an antibody specific for the Gob-5 protein, medical and
CC	diagnostic compositions containing the antisense nucleotide or the
CC	antibody, and methods and kits for screening for compounds which inhibit
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC	The antisense oligonucleotides and antibody are therefore useful for the
CC	treatment and prevention of bronchial asthma and chronic obstructive
CC	pulmonary disease. The present sequence represents a mouse Gob-5 cDNA
CC	hybridisation probe used in an exemplification of the invention.
XX	
SQ	Sequence 486 BP; 142 A; 121 C; 118 G; 105 T; 0 other;
Query Match	75.1%; Score 176.4; DB 22; Length 486;
Best Local Similarity	84.6%; Pred. No. 3, 1e-50;
Matches 198; Conservative	0; Mismatches 36; Indels 0; Gaps 0
OY	2 ACACAGCAAAATTCCTCCAGCCTGTGGATTGATTTGCATAATATGCCAAGACCTCCC 61
Db	
	197 acacaggaataattcccccagcctgttaacagtgtatgcagaacttgcgcaaggacctgcg 256
OY	62 CAATTCTCAGGGCCAGTGTACACAGCCCTTATTGAATCATCAGTGGAATGAAAACAGTTACT 121
Db	
	237 ctattctcagggccagcgctcacagccttgattgaaatctgtgaaatgaaaaaacagtaaacc 316
OY	122 TGAACTACTGATATGATGAGCAGCGTGTGATCTCTACTTAAGATGACGCGTCTACTCAA 181
Db	
	317 tggaaattactgtataacgagagcagtgcccgatgccaccaagaattgtagtgcctactcaa 376
OY	182 GGATTTTCACAACCTATATGACACGAATGATGATATACAGTGTAAAGAATGCGGGGCTC 235
Db	
	377 gtttttttacagcttttgtatgcaaatgtlagaatacagcgtlaaataatgagcttc 430
RESULT 11	
AAH46101	
ID	AAH46101 standard; DNA; 2739 BP.
XX	
AC	AAH46101;
XX	
DT	11-SEP-2001 (first entry)
XX	
DE	Mouse Gob-5 coding sequence, SEQ ID NO:3.
XX	
KM	Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
XX	expression inhibition; antisense therapy; gene therapy; bronchial asthma;
KW	chronic obstructive pulmonary disease; antiasthmatic; ds.
OS	
Mus sp.	
Key	Location/Qualifiers
CDS	1..2739
FT	/tag= a
FT	/partial
FT	/product= "Mouse Gob-5"
FT	/note= "No stop codon given in the specification"
XX	
PN	WO200138530-A1.
PD	
XX	31-MAY-2001.
PF	22-NOV-2000; 2000MO-JP08232.
XX	
PR	24-NOV-1999; 99JP-0333479.

[illegible]

AC AAC98067;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human colon cancer antigen nucleotide sequence SEQ ID NO:77.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnerracy;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder; ss.
XX
OS Homo sapiens.
PN MO20005351-A1.
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000MO-USO5883.
XX
PR 12-MAR-1999; 99US-0124270.
PA (HOMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Ruben SM;
DR WPI: 2000-587534/55.
DR P-PSDB; AAB53310.
XX
PT Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer -
XX
PS Claim 1; Page 523; 2104pp; English.
XX
CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytosolic, cardioactive, muscular,
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnerracy, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins
CC may also be used to prevent diseases such as neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, wounds, renal disorders, infectious
CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
CC AAC954007 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 1802 BP; 590 A; 348 C; 359 G; 485 T; 20 other;

Query Match 54.4%; Score 127.6; DB 21; Length 1802;
Best Local Similarity 71.5%; Pred. No. 1.9e-33;
Matches 168; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

```

QY      1 GAGCACGACCAATATGCCCGCCTGTGGACTTATGCATAATTTCGCCAAGAAGCGCTCC 60
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      61 CCAATTTTCAGGCGCAGTGTCACACCCCCTGATGATCATGTAATGGAATAACAGTTACC 120
Db          || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      730 cctgtccttgagccaatgtagtgccttcattcatgatcaacagaataagaaatacagaagt 789
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      121 TTGGAACTACTGAGTAATGAGACGACGGTGTGACTACTAAGCATGACGGTGTCTACTCA 180
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      790 ttggaaactttggataaatgtagtcaggcgctgattcttccaataaatagaatgacctacc 849
Db          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

[illegible]


```

Db 730 cctgtctcttgagagcccaatgtagcttctcattgaatcacagaatgagacatacagaagtt 789
Oy 121 TTGGAACTACTGATGATATGAGACAGGTCCTGATGCTACTAAGGATGACGGTGTCTACTCA 180
Db 790 ttggaacttttgataatgtagtcagcgctgattcttcaagaatgtagtctactcc 849
Oy 181 AGGTATTCACAACTTATGACACGAAATGATAGTATACAGTGTAAAGTGGGCTC 235
Db 850 aggtatttcaagcatatcacagaaatgtagatataagcttaaaagttcgggctc 904

```

Search completed: April 4, 2002, 03:18:10
 Job time: 60242 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:02:50 ; Search time 22700.8 Seconds

(without alignments)
170.780 Million cell updates/sec

Title: US-09-049-696-12

Perfect score: 235
Sequence: 1 GACACCACCAATTCCCGAG.....CAGTGTAAAGTGGGGGCTC 235

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBank: 1: gb_da: 2: gb_htg: 3: gb_in: 4: gb_cm: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sta: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_da: 16: em_fun: 17: em_hum: 18: em_in: 19: em_om: 20: em_or: 21: em_ov: 22: em_pat: 23: em_ph: 24: em_pl: 25: em_ro: 26: em_sta: 27: em_sy: 28: em_un: 29: em_vl: 30: em_hugo_hum: 31: em_hugo_inv: 32: em_hugo_rtd: 33: em_hug_hum: 34: em_hug_inv: 35: em_hug_rtd: 36: em_hug_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	235	100.0	2022	9 AK024970	AK024970 Homo sapi
2	235	100.0	2826	9 AF127036	AF127036 Homo sapi
3	235	100.0	3311	6 AX193489	AX193489 Sequence
4	235	100.0	3311	9 AF039400	AF039400 Homo sapi
5	177.4	75.5	3079	4 AF095584	AF095584 Sus scrofa
6	148.2	63.1	2937	10 AB017156	AB017156 Mus muscu
7	148.2	63.1	35278	9 AF039401	AF039401 Homo sapi
8	148.2	63.1	164891	2 HS651E10	HS651E10 Homo sapi
9	148.2	54.4	1895	9 AK000138	AK000138 Homo sapi
10	127.8	54.4	3204	9 AF127035	AF127035 Homo sapi
11	127.8	54.4	3204	9 AK000072	AK000072 Homo sapi
12	127.8	54.4	3204	9 AF001263	AF001263 Bos tauru
13	127.8	54.4	3204	9 AF001262	AF001262 Bos tauru
14	116.4	49.5	3022	10 AF047838	AF047838 Mus muscu
15	114.8	48.9	3471	10 AF052746	AF052746 Mus muscu
16	113.2	48.2	2765	10 AF115852	AF115852 Mus muscu
17	113.2	48.2	3137	10 BC008147	BC008147 Mus muscu
18	111.6	47.5	3058	10 AF108501	AF108501 Mus muscu
19	107.4	45.7	2832	9 AF127980	AF127980 Homo sapi
20	107.4	45.7	2970	9 AF043977	AF043977 Homo sapi
21	107.4	45.7	3604	9 AB026833	AB026833 Homo sapi
22	107.4	45.7	4077	6 AX054697	AX054697 Sequence
23	105.4	44.9	3415	9 AF043976	AF043976 Homo sapi
24	103.4	44.0	2820	4 AF001263	AF001263 Bos tauru
25	103.4	44.0	3288	4 AF001262	AF001262 Bos tauru
26	103.4	44.0	3317	4 AF001261	AF001261 Bos tauru
27	101.4	43.1	2984	4 BT036445	BT036445 Bos tauru
28	76.6	32.6	140718	2 AL356270	AL356270 Homo sapi
29	65.4	27.8	124096	2 AL354988	AL354988 Homo sapi
30	65	27.7	878	6 195746	Sequence 8
31	57.2	24.3	140718	2 AL356270	AL356270 Homo sapi
32	57.2	24.3	175591	2 AC068071	AC068071 Homo sapi
33	57.2	15.8	4042	1 AF259266	AF259266 Haemophil
34	37.2	15.1	3163	1 AF022910	AF022910 Haemophil
35	35.6	15.1	10980	1 U32749	U32749 Haemophilus
36	35.4	15.1	29605	8 AB020747	AB020747 Arabidops
37	35	14.9	80601	9 AL591472	AL591472 Human DNA
38	34.8	14.8	319000	1 AP000006	AP000006 Pyrococcus
39	34.6	14.7	108836	9 AC069137	AC069137 Homo sapi
40	34.4	14.6	39084	3 U28943	U28943 Caenorhabdit
41	34.2	14.6	339	1 HP049415	HP049415 Helicobac
42	34.2	14.6	178285	2 AC011790	AC011790 Homo sapi
43	34	14.5	167599	2 AF261720	AF261720 Homo sapi
44	34	14.5	170412	2 AC087769	AC087769 Homo sapi
45	34	14.5	173561	2 AC016240	AC016240 Homo sapi

ALIGNMENTS

RESULT	1	2
AK024970	2022 bp	29-SEP-2000
LOCUS	AK024970	PR1
DEFINITION	Homo sapiens cDNA: FLJ21317 fls, clone COJ02275, highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CACCL1) mRNA.	
ACCESSION	AK024970	
VERSION	AK024970.1	GI:10437397
KEYWORDS	oligo capping; fls (full insert sequence).	
SOURCE	Homo sapiens colon cDNA to mRNA, clone_lhb:COL clone:COL02275.	
ORGANISM	Homo sapiens	
REFERENCE	1 (sites)	
AUTHORS	Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Odayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.	
TITLE	NEDO human cDNA sequencing project	
JOURNAL	Unpublished (2000)	
REFERENCE	2 (bases 1 to 2022)	

FEATURES	source	Location/Qualifiers
COMMENT		NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' & 3' end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
JOURNAL		Sugano, S., Suzuki, Y., Ota, T., Okeyazaki, M., Nishii, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y. Direct Submission Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan (E-mail: cphn@elms.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
FEATURES		1..2022 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="COL02275" /clone_11b="COL" /tissue="type="colon" /note="cloning vector pME18SFL3"
misc_feature		1..2022 /note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CACCL) mRNA
BASE COUNT	612 a 472 c 453 g 485 t	
ORIGIN		
Query Match	100.0%; Score 235; DB 9; Length 2022;	
Best Local Similarity	100.0%; Pred. No. 1.6e-58;	
Matches 235; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1	GACACGACGAATTCCTCCAGCCCTCTGGAGATTATGCAATATTCGCCAAGAGCCTCC 60
Db	1011	GACACGACGAATTCCTCCAGCCCTCTGGAGATTATGCAATATTCGCCAAGAGCCTCC 1070
OY	61	CCAATTCTCAGGGCCAGTGTCTACAGCCCTGATTGATCAGTAAATGGAAGAAACAGTTACC 120
Db	1071	CCAATTCTCAGGGCCAGTGTCTACAGCCCTGATTGATCAGTAAATGGAAGAAACAGTTACC 1130
OY	121	TTGGAACACTACTGGATTAATGAGCAGAGTGTCTATGCTACTAAGAGATGACGGTCTTACTCA 180
Db	1131	TTGGAACACTACTGGATTAATGAGCAGAGTGTCTATGCTACTAAGAGATGACGGTCTTACTCA 1190
OY	181	AGGTAATTCACAACTTATGACACGAATGGTATATACAGTAAAGTGGGGCTC 235
Db	1191	AGGTAATTCACAACTTATGACACGAATGGTATATACAGTAAAGTGGGGCTC 1245
RESULT	2	
AF127036	2826 bp	mRNA PRI 10-AUG-1999
LOCUS		Homo sapiens calcium-activated chloride channel protein 1 (CACCL)
DEFINITION		mRNA, complete cds.
ACCESSION	AF127036	
VERSION	AF127036.1	GI:4585468
KEYWORDS		human.
SOURCE		Homo sapiens
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE		1 (bases 1 to 2826)
AUTHORS		Aguel, M., Vermet, T. and Colonscou, J. M.
TITLE		Identification of three novel members of the calcium-dependent chloride channel (CACCL) family predominantly expressed in the digestive tract and trachea
JOURNAL		FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE		99364503
PUBMED		10437792
REFERENCE		2 (bases 1 to 2826)

AUTHORS	Aguel,M. and Culouscou,J.-M.
TITLE	Direct Submission
JOURNAL	Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carriers, Rueil-Malmaison 92500, France
FEATURES	Location/Qualifiers
source	1..2826
gene	/organism="Homo sapiens" /db_xref="taxon:9606" /tissue_type="small intestine; colon"
CDS	1..2826 5..2749 /gene="CACCL1" /note="bovine epithelial chloride channel homolog" /codon_start=1 /product="calcium-activated chloride channel protein 1" /protein_id="AAD25487.1" /db_xref="GI:4585469"
translation	"MGPPKSSVFLLHLLEGALSNLQLNNNGEIVADIPNPF EDFELLQIKDMVTQASLYLEATGKRFRFNVALILPETWKTADVVRPLEYTKNA DLVAESSTPCNDDEPYTEOMCNGCEGRHLHPDFLAGKLAEYGDPRAFLHEMAH LRMGVPEBYNNDEKFYLNSNGIQQVRCSEKRLITGNVYAKCGGGCYTRKCFNKVTGJ YEKCFFVLORKEKASIMPAQVADSIVECTEOHNKKSPKNONOKCNLRSMWEVY RDSDEFKTPMTPTOPNPFSIIQIGRIYCIYLDINSGMATGNRNRLNQACLF LOYTELSGMCMATFPDAAHQSELIQINSDDDTLAKRLPAAISGRTSCSLRSR FYTRRKRPIDGESIVLLIDGEDNTISCENFEVQSGLIHFTVALGPSAOMELETLSR MTGGLQRYVASDOVNNGLIIDAFGLISSGNCASVSQSIQLESGLTLONSOMNGTVIA DSYGDKTLPLITWTQTPQPIILMDPSQKGCVNVKNTMAVLIQIGIAKVTKWKX SLOSOTLILTVTSRASNAATLPILTVYSKTNKTSKPSPLVAYANROGASVILRR SVTILSVNGKITVTELIDNGAGADAKDGVSRYETTYDTONGRYSVKRALGVYN AARARIPOOSGALYIEGWILENDIQNNPRPELNKDVOHKVCFSRTSSGSFVAS DVNPAPLPDLPEPQIIDLKAEIHGSLINTWPADPDGDGHANHYLIRISTSLDL LRDFNBSLGQVNTALLPKENASEVELFKRENTFENGDIPLAIQAADVVDLKSER SNLARVSLFIPTQPPTPEPSPDETSAACPNIHINSTIPGIHLTKMKWIGEQLSLIA"
BASE COUNT	875 a 623 c 632 g 696 t
ORIGIN	
Query Match	100.0%; Score 235; DB 9; Length 2826;
Best local Similarity	100.0%; Pred. No. 1,6e-58;
Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
D1	GACACCAGCAAAATTCGCCAGCCCTGTGGTAGTTATGCAAAATTTCGCCAAGAGACCTCC 60
D2	GACACACAGCAAAATTCGCCAGCCCTGTGGTAGTTATGCAAAATTTCGCCAAGAGACCTCC 1861
D3	CCAATTCGACGAGGCAGTGTCACACCCCTCGTAGTTATGCAAAATTTCGCCAAGAGACCTCC 1921
D4	TTTGACAAGACTGGAATAAGAGACAGGTCTGTGCTACTTAAGAGATGACGGTGTCTACTCA 180
D5	TTTGACAAGACTGGAATAAGAGAGACAGGTCTGTGCTACTTAAGAGATGACGGTGTCTACTCA 1981
D6	AGGTAATTCACACACTTATGACACGAGATGTAACAGTGTAAAGAATGTCGGGGCTC 235
D7	AGGTAATTCACACACTTATGACACGAGATGTAACAGTGTAAAGAATGTCGGGGCTC 2036
LOCUS	AX193489 3311 bp DNA PAT 15-AUG-2001
DEFINITION	Sequence 1056 from Patent W00149716.
ACCESSION	AX193489
VERSION	AX193489.1 GI:15211440
KEYWORDS	human.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 3311)
REFERENCE	Xu,J.,Lodes,M.J., Secrist,H., Benson,D.R., Neagher,M.J., AUTHORS

gene	db_xref=taxon:9623 /tissue_type="ileal mucosa"
1. .3079	
/gene="AEC"	
133. .2886	

```

BASE COUNT      892 a      758 c      755 g      674 t
ORIGIN
/funcion="involved in exocrine chloride conductance"
/codon_start=1
/product="epithelial chloride channel protein"
/protein_id="AA00077.1"
/db_xref="gi:6002646"
/translation="MGSFYRSLFLVLLHLEGAQSNSLIQNGNGSEGIYAIDPNV
EDELILQIKDMVTAKSPYLEATEERKRYEYFNVAIPASKAAPEYKELIYKNA
DVVVEEPNEENBEKFTYLSNKEQOYICSAIRGTNLPOCGSGSCVTRPGCADRYTGI
LRMVFENYENBEKFTYLSNKEQOYICSAIRGTNLPOCGSGSCVTRPGCADRYTGI
FQKCEFTLPDPOQSEKASIMPAQSIDIVVEVECKRNNKEPNQONCKIRASTWEYI
QDSDEFKRTTMTVPAPRPSLIQIQIRCAVLADLSGSMVTGGRIKLQAOAKRLT
LQPEQOAGMGMVAFDSAAAYKSELVOINSRGAEDALARSAGSTSTCSLSRASA
FTYLLKKRYPDGEELVILLDGEDNLTSGCPPEYRONAIIHTVALGSAAELELSQI
MGTGLQIYASDAQEANNQGLIDAFGALISGNRASAQSRVSLIHTJQLINEMNETV
DSYQKDTLELITERKFLSIPTFFGYPGSGRSQDSFLYKGNKKATYQVPESTARVGM
KYSLQASQTLTVSSRRSSATLPAGVATYSKMDGKPEPSPVYTKTHQGLPIIL
RAKTALEISENGKTVTLELDNAGATLNDNDISYKFPAYVANSRYSKVAALG
VNFRRRAPPLSGMCAMYIRGWIENGEIKWNPREDIKNDLOGOCYFSRASGSFV
ASVDPKSPIDPLPPCKITDILKAGIOGNILINLTWAPAGDIDHGADRYIIRSTND
LIDLDKFNDSQVMTDILPEKANESEFVYKPGIPFTNTDLFLAVAOADKTNLKS
EISNIAQVSLFLEPAPETPEPTPAPSLPCELIQVNSTITGHIILIKMKMLGELQ
SIA"

```

Query Match	75.5%	Score 177.4	DB 4	Length 3079
Best Local Similarity	84.7%	Pred. No. 1.3e-41		
Matches 199	Conservative 0	Mismatches 36	Indels 0	Gaps 0

[illegible]

RESULT	6
AB017156	
LOCUS	AB017156 2937 bp mRNA ROD 10-NOV-1999
DEFINITION	Mus musculus gob-5 mRNA, complete cds.
ACCESSION	AB017156
VERSION	AB017156.1 GI:3721911
KEYWORDS	Gob-5.
SOURCE	Mus musculus adult intestine goblet cell cDNA to mRNA.
ORGANISM	Mus musculus
REFERENCE	Eumakrolta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (sites)
AUTHORS	Komiyama, T., Tanigawa, Y. and Hirohashi, S.
TITLE	Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice
JOURNAL	Intest. Res. Commun. 255 (2), 347-351 (1999)
MEDLINE	Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)
REFERENCE	2 (bases 1 to 2937)
AUTHORS	Komiyama, T., Tanigawa, Y. and Hirohashi, S.
TITLE	Direct Submission

JOURNAL
Submitted (26-AUG-1998) to the DDBJ/EMBL/GenBank databases. Tohtu
Komiyu, ERAIO, JST, Genome Asymmetry Group, Doi Biosymmetry
Project; Tokodai 5-9-9, Tsukuba, Ibaragi 300-2635, Japan
(E-mail: komeblu.eraio.ttc-net.co.jp, Tel: 81-298-48-1515, Fax: 81-
298-47-8501)

```

FEATURES
SOURCE
1. .2937
location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="globlet cell"
/dev_stage="adult"
/tissue_type="intestine"
15. .2756
/feature="gob-5"
15. .2756
/feature="gob-5"
/codon_start=1
/protein_id="BAA33743.1"
/db_xref="GI:3721912"
translation="MESLKSPLVFLLIHLLEGVLSSELIQLNNNGEGIVIAIDHDVIV
EDLELIHKKDMVTOASPYLFEATGKRREYFNVAIILPESSKAREYRPLLEPRKNR
DVLESTSPGNDPEYTHGIGAGCGRIHLITDFDLGAKLTGLOPDDPRFVEHYMR
FRMVEFMYNNDKDEFTYLSKGRPOAVRGAALITGNNOVROCGSGITNGKVIYPRR
LYKNCVFPDPHONERKASIMFNONINSYVFCTBKHNQAPADONORCINLRSTWE
LQSEBDEKOTTPMAQPAFPFSLIGQRIVCGLVLDKSGSMLNDRIRNNNOQRRL
LQVDEGSGWAGMTFPSAAVYSELKQLNSGARDLLIKHLPIVYAGSTISGSLRPLR
AFYVIRKKRYDGEIYLTLTDEGDNLTSSCPDLVQSGAILHTVALGPAARKEQLS
KMTGLQTSYSDQVONNGVLVDPAFALSSGNALIKOHSIQLSRGVNLNNOMMNSYV
YDSVSGDITLFLIMTTHPTLFTIMDPGSGVONELITDTTKVAILQVGTAKVGFWM
YSIOASQOTLUTLYTSSRAASPTLPITVTPGVNNKGTGFPSPVYVYASIRGASIL
ASVAILTESVNGKIVTLELDLNGAGADATKNDGYISREPTAFDANGRVSXITMALGQ
TSDSORAPAPKNRMYATIDGWTDEGEVRNPNRPETSYVDOKLCEPSRSGSGSPAT
VPAAPLPDLEFPQCITDKRASIOGQNLVNLTWIARSGDYDHRASNTITKMSITIV
LNRHNSVLENTNIGLLPEKASSEIIEFELGNGTFGNGTDFIALIQADVSNLSESE
SLRAVSVFVIPAQSPEDSTPCCPDSTINSTIPGIHYVLRIMKMKLGEOMVYTLGLH
BASE COUNT
860 a 718 c 933 g 666 t
ORIGIN

```

Query Match	75.1%	Score 176.4	DB 10	Length 2937
Best Local Similarity	84.6%	Pred. NO. 2.5e-41		
Matches 198	Conservative	0	Mismatches 36	Indels 0
				Gaps 0
QY 12	ACACGACGCAAAATTCGCCAGCCCTGTGTAGTTATGCAAAATATTCGCCAAGAGCCTGCC	61		
Db 1816	ACACGAGGAATTTCCCGCAGCCCTGTATACAGTATATGCAAGCATTTCGCCAAGAGCCTGCC	1875		
QY 102	CAATTCACAGGCGCCAGTGTACACAGCCGTCGATTGAATCAAGTGAATGAAAAACAGTTACCT	121		
Db 1876	CTAATTCACAGGCGCCAGCCTCACAGCCGTCGATTGAATCTGTGAATGAAAAACAGTAACCC	1935		
QY 122	TGGAACATCTGTGATTAATGAGCAGAGGTGCTGATNGCTACHTAAGATATACGCTGTCTACTCA	181		
Db 1936	TGGAATTAATCTGGATTAACGAGCAGAGGTGCTGATNGCTACHTAAGATATACGCTGTCTACTCA	1995		
QY 182	GGATATTCACACAATATATGACACGAATGTGTAGATACAGTGTAAAGTGGCGGCTC	235		
Db 1996	GGTTTTTTACAGCTTTTGTATGCAATGTGTAGATACAGCGTTAAATATATGGGCTC	2049		
RESULT 7				
AF039401				
LOCUS				
DEFINITION	Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene,			
ACCESSION	AF039401	35278 bp	DNA	
VERSION	AF039401			
KEYWORDS	AF039401.1	GI:4009459		
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 35278)			

requests: ciomer@requests.sanger.ac.uk
On Jan 26, 2001 this sequence replaced gi:9988471

```

signature , 84004. .9135/
/note="assembly_fragment:01292
fragment chain:1"

```


CDS

/clone_1lb="COL."
/tissue_type="colon"
/note-"Cloning vector pME18SF13"
66..1475
/note-"unnamed protein product"
/codon_start=1
/protein_id="AA09069.1"
/db_xref="GI:7020031"

BASE COUNT 658 a 352 c 357 g 528 t

ORIGIN

Query Match 54.4% Score 127.8; DB 9; Length 1895;
Best local similarity 71.5%; Pred. No. 4,8e-27;
Matches 168; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

DQ 528 GACGTAAGCACTTCCCGCAGCCCTGTGGTAGTTATGCAATAATTGCCAAGACACTTC 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dq 528 GACGTAAGCACTTCCCGCAGCCCATGATGTTTGACGGAATCTCAAGAAGATATGTA 587

OY 61 CCAATTCACAGCCGCAGCTGTACACACCCTGATTTGATTAATGATGATGAAAACACTTACC 120
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 598 CCTGTCTTGGACAAGATGTAAGTACTGCTTTCAATGATACAGAAAGACATACAGAACTT 647

OY 121 TTGGAAGTACAGTAATGAGACAGCTGTGATGCTACTAAGATGAGACGCTGTACTTCA 180
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 TTGGAAGTCTTTGGATTAATGCTGACAGGCCCTCTTCTTCCAAGATGATGAGCTACTTCC 707

OY 191 AGGTATTTACAACTTTATGACACGAATGTAGATACACTGTAATAAGTCGGGGCTC 235
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 708 AGGTATTTTACAGCATATATACAGAAATGGCAGATTTACTTTAAAGTGGGGCTC 762

RESULT 11
AF127035

LOCUS AF127035 PRI 11-AUG-1999

DEFINITION Homo sapiens calcium-activated chloride channel protein 2 (CACCC2)

ACCESSION AF127035 mRNA complete cds.

VERSION AF127035

KEYWORDS AF127035.1 GI:5726288

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3204)
Aguel,M., Vernet,T. and Culousscou,J.-M.
Identification of three novel members of the calcium-dependent
chloride channel (CACCC) family predominantly expressed in the
digestive tract and trachea
FEBS Lett. 455 (3), 295-301 (1999)

JOURNAL MEDLINE
PUBMED 99364503
10437792
2 (bases 1 to 3204)
Aguel,M. and Culousscou,J.-M.
Direct Submission
Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des
Carrieres, Neuilly-Malmaison 92500, France
location/Qualifiers
1..3204
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="colon"
1..3204
gene

CDS

```

/gene="CACCC2"
/29..2782
/gene="CACCC2"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 2"
/protein_id="AAD8398.1"
/db_xref="GI:5726289"

```

BASE COUNT	1098 a	594 c	633 g	879 t
ORIGIN				

Query Match	54.48;	Score 127.8;	DB 9;	Length 3204;
Best Local Similarity	71.58;	Pred. No. 4.7e-27;		
Matches 168;	Conservative	0;	Mismatches 67;	Indels 0;
				Gaps 0;

OY	1	GACACCCAGCAAATTCCTCCACCCTCGTGTGATTATGCACAATAATGCCAACGAGCCTCC	60
Db	1835	GAGCTAACAACGTTTCCCACCCCAATGATTGTTTACCGCAAAATTTCTACAGAGATATGTA	189
OY	61	CCAATTCCTCAGGGCCACTGTGCACAGCCCTGATTGAATCAGTGATGAAAAACAGTTACC	120
Db	1895	CCGTCTCTGGAGCCCATGTGACGTGGCTTCATTGAATCACAGAATGAGACATACAGAGTT	195
OY	121	TTCGAACTACTGATTAATGAGCAGCTGCTGATGCTACTAAGAGTAGCGTGTCTACTCA	180
Db	1955	TTGGAACTTTTGGATGTATGTGTGACAGGGCGCTGATCTTTCAAGATGATGTGAGCTCACTCC	204
OY	181	AGGATATTCACAATCTATGACAGCAAGATGAGATACACGTGTAAAATCGGGCTC	235
Db	2015	AGGATATTTTACAGATATPACAGAAATGGGAGATATNAGCTTAAAAATTTGGGGCTC	2069

RESULT	12
AK000072	
LOCUS	
DEFINITION	AK000072 3221 bp mRNA
ACCESSION	AK000072
VERSION	AK000072.1 GI:7019922
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens colon CDNA to mRNA, clone_11b:COL clone:COL01613.
ORGANISM	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (stiles)
AUTHORS	Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Ohtani,R., Oka,W., Suzuki,Y., Ohashi,M., Nishi,T., Shihabara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished (2000)
REFERENCE	2 (bases 1 to 3221)
AUTHORS	Sugano,S., Suzuki,Y., Ota,T., Ohashi,M., Nishi,T., Isogai,T., Shihabara,T., Tanaka,T. and Nakamura,Y.
TITLE	Direct Submission
JOURNAL	Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sutils

COMMENTARY
Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency). *Received 20/11/2000*

FEATURES

Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01613"
/clone_1lb="COL"
/tissue_type="colon"
/mode="cloning vector pME185FL3
1..3221"

```

BASE COUNT	1105 a	600 c	634 g	882 t
ORIGIN				

Query Match	54.48;	Score 127.8;	DB 9;	Length 3222;
Best Local Similarity	71.58;	Pred. No. 4.7e-27;		
Matches 168;	Conservative 0;	Mismatches 67;	Indels 0;	Gaps 0;

Q7	1	GACACAGCAAAATTC	CCCCACACCTCTG	AGTTATGCAAAATATTC	GGCAAGGACCTCC	60
D5	18:19	GAGCTTAACAGTTT	CCCCACGCCCAATG	ATGTTTATGCGCAAAATCT	CACAGATATGTA	19080
QY	91	CCAAATTCACAGGGC	CACTGTGCACAGCCCT	GATTCAGTACAGTAAAG	AAAAACAGTTACC	120
D5	19:9	CCCTTCTTGGAGC	CAATGTGACTGCTTTC	ATTGATGATACAGAAATG	AGACATACAGACTT	19680
QY	121	TTGGAACACTCGAT	TAATGAGCAGCAGGT	CTGTGATGCTACTA	AGAGTAGCGGTCT	180
D5	19:9	TTGGACATTTT	TGATATGCTGCAGCG	CGCTGATCTTTCA	CAAGATATGATG	20280
QY	141	AGGATATTCACAA	CTTGTGACAGATNG	ATATACAGTTAAAGT	GGGGGCTC	235
D5	20:3	AGGATATTTAC	AGCATATACAGAAAT	ATGACAGATATAGCT	TATTAACCTTGGGCTC	2083

	RESULT	13						PAT	21-MAR-2001
AX092338	LOCUS	AX092338							
DEFINITION	Sequence 69 from Patent WO0116318,								
ACCESSION	AX092338								
VERSION:	AX092338.1								
	GI:13444483								

BASE COUNT	1159 a	596 c	632 g	878 t
ORIGIN	.			

RESULT	14
AF047838	
LOCUS	3022 bp mRNA
DEFINITION	Mus musculus calcium-sensitive chloride conductance protein-1
ACCESSION	(MCCAC) mRNA, complete cds.
VERSION	AF047838
KEYWORDS	AF047838.1 GI:3925280
SOURCE	house mouse.

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	TITLE	FEATURES	source	gene	CDS
Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Gandhi, R., Ebble, R.C., Gruber, A.D., Schneur, K.D., Ji, H.L., Fuller, C.M. and Pauli, B.U.	Molecular and functional characterization of a calcium-sensitive chloride channel from mouse lung	J. Biol. Chem. 273 (48), 32096-32101 (1998)	99041980				2 (bases 1 to 3022)	Gandhi, R., Ebble, R.C., Gruber, A.D. and Pauli, B.U.		
									Direct Submission	Submitted (12-FEB-1998) Vet. Pathology, Cornell Univ., VMC C4-146, Ithaca, NY 14856, USA		
									Location/Qualifiers			
									1..3022			
									/organism="Mus musculus"			
									/db_xref="taxon:10090"			
									/tissue_type="whole lung"			
									1..3022			
									/gene="mclcal1"			
									18..2726			
									/note="mclcal1"			
									/note="calcium-activated chloride channel; Alternative cyclic channel to CFTR that may influence severity of Cystic Fibrosis in different tissues; mCAC-1"			
									/codon_start=1			
									/product="calcium-sensitive chloride conductance protein-1"			
									/protein_id="AAC79982.1"			
									/db_xref="GI:3925281"			
									/translation="MWPGIOVLRTYLTHLLINTESSMYHLNSNGEGVYLAINSPVDERLIPKREKNTQASTYLEFASQGRVYFRNLSILVPTWRSKEPTLPKRESDKAKVIVADPHLGDDDPYTLQYGGCGDGOIHTFPNLTLDNLRIYGRPRVYHWAHIRMGEFENVDNDPPMYSRKNTLEARTGSTRITVWVNCERGNCTFACRDSRTIRYEPCEFTIPDKIQTAGASIMPMQNLNSVFECTENHNAEAPLNQKCNRSRTWDVKTSDPONAPNRGTETAPRPPTSLKRSRRVVCVLDKSGMDEKRLIRNNOAEELVYIOIVKESWVGVLTFPDSAHIONYLKIKRSSDYOKITANLPQOAGGTSICIGDILAGPATISSDOSTSGSEYVLTLDGEDNGIRCFEFAVNSGALIIITIALGSPARLELELSDMGTRFTYANKDNLSDAFSIRISITDGSGVYTSDFODDKLIRNSALQIPGTAETVLDSTVGNDTFEYVWVKKPEIILQDPDGKRYTSDFODDKLIRNSALQIPGTAETV			

Query Match	Score	DB	Length
Base Local Similarity	69.1%	Pred. No. 1e-23;	
Matches 159;	Conservative	0;	Mismatches 71; Indels 0; Gaps 0;
3	CACAGCGAAATTCOCACGACCTGTGTAGTATATGCAAAATFTGGCCAAAGAGCOTCCCC	62	
DB	1892	CACAGCCAGTACCCGTACCCGGATGATGTAGCCACAGGGCGACCCAGAGATTTTTCG	1891
QY	93	AATTCACAGGCCAGGTGCACAGCCCTGATGTAATCACTGTAATGGAAGAAAAACATTTACCTT	122
DB	1892	TGTTCTGGGAGCCAAATGTCACAGCCCTCATAGAAAGCTGAACATGSAACATCAAGTACCTT	1951
QY	123	GGAACTACTGGATAATGAGCAGGTGCTGATGATCTACTAAGATGACGCTGTCTACTCAAG	182
DB	1932	GGAGCTCGGACAAATGGGGCAGGTGCTGATCTGTTAAAAATGATGGCATCTACACAAAG	2011
QY	183	GATTTTCACCACTATGACAGCAATGGATGATGACAGTGTAAAGAGCGGGG	232
DB	2012	ATACTTTACCAATATCTCATGGAATAGTGATGACAGCTTAAAGAGTGGTG	2061

FEATURES	Source	LOCUS	DEFINITION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE
AF052746	3471 bp	mRNA	Mus musculus chloride channel Cacc mRNA, complete cds.	AF052746	GI:3560546	house mouse.	1 (bases 1 to 3471)	Romio,L., Musante,L., Cinti,R., Serl,M., Moran,O., Zegarra-Moran,O.	Characterization of a murine gene homologous to the bovine Cacc chloride channel	Gene	228 (1-2), 181-188 (1999)	2 (bases 1 to 3471)	Romio,L., Musante,L., Cinti,R., Moran,O., Serl,M. and Galletta,L.J.V.	Direct Submission	Submitted (09-MAR-1998) lab. di Genetica Molecolare, Istituto Giannina Gaslini, Genova I-16148, Italy	AF052746
AF052746	3471 bp	mRNA	Mus musculus chloride channel Cacc mRNA, complete cds.	AF052746	GI:3560546	house mouse.	1 (bases 1 to 3471)	Romio,L., Musante,L., Cinti,R., Serl,M., Moran,O., Zegarra-Moran,O.	Characterization of a murine gene homologous to the bovine Cacc chloride channel	Gene	228 (1-2), 181-188 (1999)	2 (bases 1 to 3471)	Romio,L., Musante,L., Cinti,R., Moran,O., Serl,M. and Galletta,L.J.V.	Direct Submission	Submitted (09-MAR-1998) lab. di Genetica Molecolare, Istituto Giannina Gaslini, Genova I-16148, Italy	AF052746

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:16 ; Search time 16681 Seconds

(without alignments)
141.722 Million cell updates/sec

Title: us-09-049-696-5

Perfect score: 220
Sequence: 1 CTATAGTGAATTCGTACA.....GACAAATCTGCAACATGGC 220

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: qb_est1:*
11: qb_est2:*
12: qb_hlc:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_tod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	220	100.0	480	10	AA581198 nd38c07.r
2	219	99.5	460	10	AA750670 CM2-CN03
3	200.8	91.3	307	10	AA577823 CM0-CN000
4	162	73.6	2915	12	AA007466 Mus muscu
5	162	73.6	2933	12	AA008659 Mus muscu
6	138.6	63.0	563	13	AA0437882 HS_5070_A
7	138.4	62.9	890	11	AA498535 vH40F12.r
8	125.4	57.0	427	10	AA498535 vH40F12.r
9	114.2	51.9	435	10	AA844161 RC4-CN04
10	107.2	48.7	395	13	AA2465176 1M0274L23
11	102.2	46.5	949	11	AA099712 601751347
12	102.2	46.5	972	11	AA0173173 602335553

Result No.	Score	Query Match	Length	DB ID	Description
13	100.8	45.8	744	13	AA855898 2M0160E12
14	99	45.0	402	10	AA656570 UI-M-BHO-
15	99	45.0	640	11	AA320704 uz55f02.y
16	98.4	44.7	935	11	AA099600 601751621
17	97.6	44.4	650	13	AA035021 BH035021
18	97.4	44.3	305	10	AA240298 BB240298
19	91.6	41.6	507	13	AA203867 RPT-23-3
20	90.8	41.3	547	10	AA380745 601270855
21	88.4	40.2	1047	11	AA247065 602359723
22	82.4	37.5	602	10	AA602490 DKEP68P
23	82.4	37.5	753	11	AA598508 602631026
24	80.8	36.7	880	11	AA598508 DKEP313M
25	79.2	36.0	580	11	AA42391 370140 MA
26	73.4	33.4	392	10	AA677576 602624129
27	58.2	26.5	372	11	AA755896 v04f06.r
28	56	25.5	858	11	AA110624 602896268
29	54.2	24.6	1016	11	AA102225 601752527
30	45	20.5	637	10	AA750773 CM0-CN004
31	41.2	18.7	704	11	AA6565023 602833653
32	38.6	17.5	696	10	AA564941 601343653
33	38.4	17.5	584	10	AA514035 AL514035
34	38	17.3	758	11	AA401353 602465468
35	38	17.3	922	13	AA0073W CNS0073W
36	37.4	17.0	688	13	AA066784 Drosophila
37	37.2	16.9	678	13	AA174286 Tetradon
38	37	16.8	807	11	AA720797 RPT-24-8
39	36.8	16.7	1104	11	AA684530 602140838
40	36.6	16.6	118	10	AA179009 602330714
41	36.6	16.6	118	10	AA841718 RC3-CN001
42	36.2	16.5	884	13	AA065523 Drosophila
43	36	16.4	792	13	AA100446 RPT-24-3
44	36	16.4	430	13	AA584450 n005901.s
45	36	16.4	530	13	AA211088 RPT-24-1
			926	13	AA0752761 HS_5569_B

ALIGNMENTS

RESULT 1
AA581198 480 bp mRNA EST 05-JAN-1998
nd38c07.r1 NCT-CCAP-001 Homo sapiens cDNA clone IMAGE:802572.5,
similar to TR:G1184066 G1184066 CALCIUM-ACTIVATED CHLORIDE CHANNEL.
; mRNA sequence.

ACCESSION AA581198
VERSION AA581198.1 GI:2358970
KEYWORDS EST.

SOURCE 1
ORGANISM human:

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 480)
NCT-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael
Emmerit-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies Inc., David Krizman,
Ph.D.

CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCT-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/dbtrp/image/image.html
Seq primer: -28m13 rev1 ET from Amersham

FEATURES
Location/Qualifiers
1..480

/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 168 a 101 c 97 g 114 t
 ORIGIN
 /clone="IMAGE:802572"
 /clone_id="NCI-GAP_C01"
 /tissue_type="bulk tumor"
 /lab_host="DH10B"
 /note="Organ: colon; Vector: pCMV-SPORT2; Site:1: SalI;
 Site:2: NotI; cloned unidirectionally. Primer: Oligo dr."

Query Match 100.0%; Score 220; DB 10; Length 480;
 Best Local Similarity 100.0%; Pred. No. 6.7e-39;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTATAGTGAATTCCTGTACAGACAAACCAAGAGCTCCCAACAGCAAAATC 60
 DB 215 CTATAGTGAATTCCTGTACAGACAAACCAAGAGCTCCCAACAGCAAAATC 274
 QY 61 AAAATGCAATCTCCGAGACATGGAGATGATCCGTATCTGAGACTTTAAGAAA 120
 DB 275 AAAATGCAATCTCCGAGACATGGAGATGATCCGTATCTGAGACTTTAAGAAA 334
 QY 121 CCACCTCTATGACACACAGCCACCAATCCACCTCTCATGCTGAGATTGACAAA 180
 DB 335 CCACCTCTATGACACACAGCCACCAATCCACCTCTCATGCTGAGATTGACAAA 394
 QY 181 GAATGTGTCTTACTCTTGACAAATCTGGAAGCATGGC 220
 DB 395 GAATGTGTCTTACTCTTGACAAATCTGGAAGCATGGC 434

RESULT 2
 AM750670 460 bp mRNA EST 28-APR-2000
 LOCUS CM2-CN0039-110100-069-c01 CN0039 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM750670
 ACCESSION AM750670.1 GI:7665602
 VERSION EST.
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 460)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.G.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL
 MEDICAL
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM2&t2=CM2-CN0039-
 110100-069-c01&t3=2000-01-11&t4=1)
 Seq primer: puc 18 forward.

FEATURES
 source
 1. 460
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CN0039"
 /dev_stage="Adult"
 /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;

Site:2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 114 a 96 c 100 g 150 t
 ORIGIN

Query Match 99.5%; Score 219; DB 10; Length 460;
 Best Local Similarity 100.0%; Pred. No. 1.1e-38;
 Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTATAGTGAATTCCTGTACAGACAAACCAAGAGCTCCCAACAGCAAAATC 60
 DB 361 CTATAGTGAATTCCTGTACAGACAAACCAAGAGCTCCCAACAGCAAAATC 302
 QY 61 AAAATGCAATCTCCGAGACATGGAGATGATCCGTATCTGAGACTTTAAGAAA 120
 DB 301 AAAATGCAATCTCCGAGACATGGAGATGATCCGTATCTGAGACTTTAAGAAA 242
 QY 121 CCACCTCTATGACACACAGCCACCAATCCACCTCTCATGCTGAGATTGACAAA 180
 DB 241 CCACCTCTATGACACACAGCCACCAATCCACCTCTCATGCTGAGATTGACAAA 182
 QY 181 GAATGTGTCTTACTCTTGACAAATCTGGAAGCATGG 219
 DB 181 GAATGTGTCTTACTCTTGACAAATCTGGAAGCATGG 143

RESULT 3
 AM577823 307 bp mRNA EST 16-MAR-2000
 LOCUS CM0-CN0004-130100-139-h03 CN0004 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM577823
 ACCESSION AM577823.1 GI:7252872
 VERSION EST.
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 307)
 HCGP <http://www.ludwig.org.br/ORESTES>.
 The FAPESP/LICR Human Cancer Genome Project
 Unpublished (1999)
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM0&t2=CM0-CN0004-
 130100-139-h03&t3=2000-01-13&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 33
 High quality sequence stop: 307.

FEATURES
 source
 1. 307
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CN0004"
 /dev_stage="Adult"
 /note="Organ: colon_normal; Vector: puc18; Site:1: SmaI;
 Site:2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under

LOCUS	AK008659	2933 bp	mRNA	HTC	05-JUL-2001
DEFINITION	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210006611, full insert sequence.				
ACCESSION	AK008659				
VERSION	AK008659.1	GI:12842987			
KEYWORDS	CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone:11b:RIKEN full-length enriched mouse cDNA library clone:2210006611.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 2933)				
TITLE	Carninci,P. and Hayashizaki,Y				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Methods in enzymology. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	10349636				
AUTHORS	2 (bases 1 to 2933)				
TITLE	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,				
JOURNAL	Itou,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
MEDLINE	Normalization and subtraction of cap-trapper-selected cDNAs to				
PUBMED	prepare full-length cDNA libraries for rapid discovery of new genes				
REFERENCE	Genome research. 10 (10), 1617-1630 (2000)				
AUTHORS	3 (bases 1 to 2933)				
TITLE	Shibatani,K., Itou,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,				
JOURNAL	Konno,H., Akiyama,Y., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,				
MEDLINE	Sunji,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,				
PUBMED	Yamanoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,				
REFERENCE	Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wachihi,M.,				
AUTHORS	Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,				
JOURNAL	Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.				
MEDLINE	RIKEN integrated sequence analysis (RISA) system-384-format				
PUBMED	sequencing pipeline with 384 multichannel sequencer				
REFERENCE	Genome research. 10 (11), 1757-1771 (2000)				
AUTHORS	4 (bases 1 to 2933)				
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the				
JOURNAL	FANTOM Consortium.				
MEDLINE	Functional annotation of a full-length mouse cDNA collection				
PUBMED	Nature 409, 685-690 (2001)				
REFERENCE	5 (bases 1 to 2933)				
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,				
JOURNAL	Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,				
MEDLINE	Harakawa,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Horii,F.,				
PUBMED	Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,				
REFERENCE	Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matuyama,T.,				
AUTHORS	Miyazaki,A., Nishi,K., Nomura,K., Nunazaki,R., Ohno,M., Okazaki,Y.,				
JOURNAL	Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,				
MEDLINE	Sasaki,D., Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T.,				
PUBMED	Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,				
REFERENCE	Tanaka,T., Tajima,Y., Toyota,T., Yamamura,T., Yasunishi,A.,				
AUTHORS	Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.				
JOURNAL	Direct Submission				
MEDLINE	Submitted (10-JUL-2000) Yoshida Hayashizaki, The Institute of				
PUBMED	Physical and Chemical Research (RIKEN), Laboratory for Genome				
REFERENCE	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),				
AUTHORS	RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama				
JOURNAL	Kanagawa 230-0045, Japan [E-mail:genome-resgsc.riken.go.jp,				
MEDLINE	URI:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,				
PUBMED	fax:81-45-503-9216)				
REFERENCE	Please visit our web site (http://genome.gsc.riken.go.jp/) for				
AUTHORS	further details.				
JOURNAL	cDNA library was prepared and sequenced in Mouse Genome				
MEDLINE	Encyclopedia Project of Genome Exploration Research Group in Riken				
PUBMED	Genomic Sciences Center and Genome Science Laboratory in RIKEN.				
REFERENCE	Division of Experimental Animal Research in Riken contributed to				
AUTHORS	prepare mouse tissues. First strand cDNA was primed with a primer				
JOURNAL	[5' GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTT 3'], cDNA was				
MEDLINE	prepared by using tetrahase thermo-activated reverse transcriptase				
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
JOURNAL					
MEDLINE					
PUBMED					

and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapted of sequence 15',

GAGGAGGAGATTCGCGAGCTTATTTAAATTAACCCCCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

FEATURES	SOURCE	Location/Qualifiers
1..2933		/organism="Mus musculus"
		/strain="C57BL/6J"
		/db_xref="taxon:10090"
		/db_xref="MGD:MGI:1346342"
		/db_xref="MGD:MGI:189583"
		/clone="221006g11"
		/sex="male"
		/tissue_type="stomach"
		/clone_lib="RKEN full-length enriched mouse cDNA library"
		/dev_stage="adult"
29..2770		/note="putative"
		/codon_start=1
		/protein_id="BAB25815.1"
		/db_xref="GI:12842988"
		/translation="MESLKSPPVLLIHLLEGVLSSESLIOLNNNGEGIVAIIDHVP EDELLOHIDKDMYQASPRYLFEATGRKRPFFKNVALLPESSKAPETRPLETFEKNNA DVLSTTSPLGNDEPTEHTHIGACGCEKGRIRHLITDPLAGKLTIOGYQDRTFVEMAH FRWVFNENDEKFEVLSKGPQVAVRCSAALTGNQVRRCGGSCITNGKCIDRVAG LYKNQCVFDPDHPONEKASIMENONINSVEFECKEHNNDAPDNQRCULNSTWE IOEEDFKQTPMAQPPAPFPFSLIQQRIVCLVDLKSGLMDRLNRNQAISLE LLOVEQSGWGWMTFDSAAYVOSELKOLNSGARDDLKILPTVASAGTSICGLRT AFVYIKRKYPTDQSEIILTDGEENITSSCDFLVKQSGAIIHTVALPAAKETEOLS KMNGEITSDSDQVNNNGIVDAFPAALSSGNAIIQHSIOLESRCGVLONNMMNGSVY VDSVCKQDTLELIMTHHPPIPIFIIMDQVQNGEFLIDTTTKVYILOYCPKAPFKPK YSDASSQTLTITTSRAASATLEPITVTPVYKNTGKFPSPVTVYSIRQASPILR ASVALIESVNGKTVTLELDNGAGADATKNDGYSRFTAFDANGRSVKIMLGGV TSDORAPVKNRNPYIDGWIJEDGVRNPNRPRTSYVDQQLCFSSSGSFPVATN VPAAPPIDLEPFCQITDLKASIOGOMLVNLTWPAQDGYDHGRASNYIIMSTSIYD LRHFNPTSLQVNTGLIIPKREASSEIEPEFELGNTFGGTDIPIAIOAVDQSNLKEE SNARVSVETIPAOGRPIPEDSTPPCPDISINSTIPGIVHLKIMKMWJEGMVTGLH
2912..2917		/note="putative"
2933		/note="putative"
		/note="putative"
polya_site		
BASE COUNT	845 a 722 c 697 g 669 t	
ORIGIN		
Query Match	73.6%	Score 162; DB 12; Length 2933;
Best Local Similarity	83.9%	Prod. No. 2.4e-26;
Matches 183; Conservative	0; Mismatches 35; Indels 0; Gaps 0;	
OY	1	CTATGTTGAATCTGTCACAGCAAAACCCAAAGAAAGACCTCCCAAAACAAGCAAAATC 60
LD	755	CTGTGTGAATTCGTACAGAAAAAATCAACAATCAAGAAAGCCCAAAATGACCAAAACC 824
OY	31	AAAAATGCAATCTCCGAGCAACATGGAAGTATCTGATTCGTAGAGACTTTAAGACAA 120
DB	835	AACGATGCAATCTCCGAGCAACATGGAAGTATCTGATTCGTAGAGACTTTAAGACAA 884
OY	131	CCATCTGATGACACAGACAGCCACCAAAATCCACCTTCTCATTTGCTGCGAGATGGACAA 180
DB	885	CCACCTCCATGACACAGACAGCCACCAAAATCCACCTTCTCATTTGCTGCGAGATGGACAA 944
OY	151	GAATGTTGTTAGTCTGTCGACCAATCTGGAAGCATG 218
DB	945	GAATGTTGTTAGTCTGTCGACCAATCTGGAAGCATG 982
RESULT	6	
AA0437882	563 bp	DNA
LOCUS		GSS 31-MAR-1999
DEFINITION	HS_5070_AL_D02_T7_RFC1-11 Human Male BAC Library Homo sapiens	

ACCESSION	genomic clone Plate=646 Col=3 Row=G, DNA sequence.					
VERSION	A0437882					
KEYWORDS	A0437882.1 GI:4549221					
SOURCE	GSS.					
ORGANISM	human.					
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 563)					
AUTHORS	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J.J., Young,J.J., Zhao,S., Adams,M.D. and Hood,L.					
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome					
JOURNAL	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)					
MEDLINE	99380589					
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones are derived from the human BAC library RPci-11. For BAC library availability, please contact Pieter de Jong (pietere@biology.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://Dacpac.med.buffalo.edu/ordering_bac.htm) or from Reserach Genetics (info@resgen.com). BAC end Web Server: http://www.hscg.washington.edu Plate: 646 row: G column: 3 Seq primer: T7 Class: BAC ends High quality sequence stop: 563. Location/Qualifiers 1..563 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="plate-646 Col-3 Row-G" /clone_lib="RPci-11 Human Male BAC Library" /sex="male" /note="Vector: pBACE3.6; Site.1: EcoRI; Site.2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"					
BASE COUNT	166 a 149 c 99 g 143 t 6 others					
ORIGIN						
Query Match	63.0%; Score 138.6; DB 13; Length 563;					
Best Local Similarity	86.8%; Pred.No. 4.8e-21;					
Matches 164; Conservative	0; Mismatches 24; Indels 1; Gaps 1;					
OY	33 AACAAAGAGCTCCAAACAGCAAAATCAAAAATGCATCTCGGAAGACA-TGGAGT 91					
Db	95 AAACCAAGACTCCAAACAAGCCAATCATATATGTCATCTCGGAAGACANTGTATT 154					
OY	92 GATCGGTATTTTGAGGACTTTAGAAAAACCAGCTCTATGACACAGCACCACAAATCC 151					
Db	155 GATCGGATTTCTGAGGACTTAACAAAACCTCTCTATGACACAGCACCACAAATCC 214					
OY	152 CACCTTCATTTGGCGAGATTTGGCAAGAATTTGTGTTTAGCTTGACCAATCTGG 211					
Db	215 CACCTTCATTTGCTGCAGATGACACAAGAATTTGTGTCAAGTCTTGACCAATCTGG 274					
OY	212 AAGCATGCC 220					
Db	275 AAGCACGGC 283					
RESULT 7						
BP578833	890 bp mRNA EST 12-DEC-2000					
LOCUS						

```

DEFINITION 6020948761 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4209091 5'
ACCESSION mRNA sequence.
VERSION BF578893
KEYWORDS BF578893.1 GI:11652545
SOURCE EST.
ORGANISM house,mouse
            Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LHM9774 row: h column: 20
            High quality sequence stop: 669.
FEATURES
            Location/Qualifiers
                1..890
                /organism="Mus musculus"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:4209091"
                /clone_lib="NCI_CGAP_Co24"
                /lab_host="DH10B (TI phage-resistant)"
                /note="Organ: colon; Vector: pCMV-SPORT6; Site.1: NCI;
                Site.2: Salt; Cloned unidirectionally. Primer: Oligo dt.
                Average insert size 1.6 kb. Constructed by Life
                Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 275 a 217 c 233 g 165 t
ORIGIN
Query Match 62.9%; Score 138.4; DB 11; Length 890;
Best Local Similarity 80.3%; Pred. No. 4.8e-21;
Matches 175; Conservative 0; Mismatches 41; Indels 2; Gaps 1
Oy 11 CTATAGTGAATTCGTGTACAGAACAAACCCAGCAACAAAGAGCTCCAAACGACAAATC 60
Db 231 CCGTGGTGAATTCGTGTACAGAAAAAATCACAATCAACGAAGAGCCCAATGACCAAAAC 290
Oy 61 AAAATGCAATTCGGAAGACACATGGGAAGTATCGGTATCTGTAGGACTTTAAGAAAA 120
Db 291 AACGATGCAATTCGGAAGACAGCTGGGAAGTATCCAGGAATCTTGAGAGCTTCAACAAA 350
Oy 121 CCATCCCTGTGACACACACAGCCACCAAAATCCACCTTCATCTGTGACGATTGACAAA 180
Db 351 CCATCCCATGACAGCCACAGCCACCACTTCGACACCACTTCACACTGCTGCAAAAT--GGACAA 408
Oy 131 GAATTTGTGTTTACTGCTTGACAAATCTGAGACATG 218
Db 409 AGAATGTGTGCTTAGTTCTTGATAGTCCGGAGCATG 446
RESULT 8
LOCUS AA498535 427 bp mRNA EST 01-JUL-1997
DEFINITION Vh04f12.1 Barstead mouse pooled organs MRLRB4 Mus musculus CDNA
clone IMAGE:889487 5' similar to TR:G1184066 G1184066
CALCIUM-ACTIVATED CHLORIDE CHANNEL. // mRNA sequence.
ACCESSION AA498535
VERSION AA498535.1 GI:2233558
KEYWORDS EST.
SOURCE house,mouse;
            Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 427)	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Unpublished (1996)		
1	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B., Weising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.	The WashU-HMMI Mouse EST Project	Washington University School of Medicine	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
				Tel: 314 286 1800
				Fax: 314 286 1810
				Email: mouseest@wustl.edu
				This clone is available royalty-free through LLNL; contact the
				IMAGE Consortium (info@image.llnl.gov) for further information.
				MG1:517447
				Seq primer: -28m13 rev2 ET from Amersham
				High quality sequence stop: 377.
				Location/Qualifiers
				1. 427
				/organism="Mus musculus"
				/strain="FVB/N"
				/db_xref="taxon:10090"
				/clone="IMAGE:889487"
				/clone_lib="Barsteed mouse pooled organs MPRB4"
				/sex="mixed"
				/tissue_type="pooled organs"
				/dev_stage="7 day"
				/lab_host="DH10B"
				/note="Organ: pooled; Vector: pTY73D-Pac (Pharmacia) with
				a modified polylinker; Site.1: EcoRI; Site.2: NotI; 1st
				strand cDNA was primed with a Not I - oligo(dT) primer [5
				TGTTACGATCTGCAAGTGGAGGCGGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
				3']; double-stranded cDNA was ligated to Eco RI adaptors
				[GTGGATCTCGTACG], digested with Not I and cloned into
				the Not I and Eco RI sites of the modified pTY73 vector.
				Library constructed by Bob Barsteed."
BASE COUNT	143 a 108 c 83 g 93 t			
ORIGIN				
Query Match	57.0%; Score 125.4; DB 10; Length 427;			
Best Local Similarity	82.3%; Pred. NO. 4e-18;			
Matches 144; Conservative	0; Mismatches 31; Indels 0; Gaps 0;			
1	CGATGTTGAATTCGTGTACGAGAAACCAACCAAGACGCTCCAAACGACAAATC 60			
68	CGTGGTGAATTCGTGTACGAGAAACCAACCAATCAGAGAGCCCAATGACCAAAAC 127			
61	AAAAATGCATCTCCGAGACACATGGGAGTAGTCGATTCGTGAGAGACTTAAAGAAA 120			
128	AACGATGCATCTCCGAGACACATGGGAGTAGTCGAGAGACTTAAAGAAA 187			
QY	121 CCAGCTCTATGACACAGACCAACCAATCCACCTTCATTCGTCGAGATTGG 175			
DB	188 CCAGCTCCATGACAGCCAGCACCTGCAGCACCACTTCCTACTGTTCGAAAAATGG 242			
RESULT 9				
LOCUS	AW844161 435 bp mRNA EST 18-MAY-2000			
DEFINITION	RC4-CN0048-150400-013-f07 CN0048 Homo sapiens cDNA, mRNA sequence.			
ACCESSION	AW844161			
VERSION	AW844161.1 GI:7938144			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE	1 (bases 1 to 435)			

FEATURES	SOURCE
1. 435	/organism="Homo sapiens"
/db_xref="taxon:9606"	
/clone_lib="cDNA048"	
/dev_stage="Adult"	
/note="Organ: colon, normal; Vector: puc18; Site: 1: Sma1; Site: 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
location/Qualifiers	
124 a	105 c 115 g 91 t
Query Match	51.9%; Score 114.2; DB 10; Length 435;
Best local Similarity	93.7%; Pred. No. 1,2e-15;
Matches 119; Conservative	0; Mismatches 8; Indels 0; Gaps 0;
QY	154 TCCGATGATTCGTGAGACTTTAAGAAAACCACTCCATGATGACAAACAGCCACCAAAATCCCA 153
DB	28 TCCGATGATTCGTGAGACTTTAAGAAAACCACTCCATGATGACAAACAGCCACCAAAATCCCA 87
QY	154 CCTTTCATTCGTGAGACTTTAAGAAAACCACTCCATGATGACAAACAGCCACCAAAATCCCA 213
DB	38 CCTTTCATTCGTGAGACTTTAAGAAAACCACTCCATGATGACAAACAGCCACCAAAATCCCA 147
QY	214 GCATGAC 220
DB	148 GCATGAC 154
RESULT 10	
AZ465176	395 bp DNA
LOCUS	IM0274123R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
DEFINITION	clone UUCG1M0274123 R, DNA sequence.
ACCESSION	AZ465176
VERSION	AZ465176.1 GI:10623301
KEYWORDS	GSS.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
AUTHORS	Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,

REFERENCE 1 (bases 1 to 972)
 AUTHORS NIH-MGC <http://mhc.mcl.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.jnl.gov>
 Plate: LHAM0257 row: 1 column: 14
 High quality sequence stop: 534.

FEATURES

source
 1. 972
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="FVR/N"
 /db_xref="taxon:10090"
 /clone_image="458709"
 /clone_lib="NCI-CGAP_Maml"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="3 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt: Site: 2; Note: Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
 BASE COUNT 284 a 278 c 208 g 202 t
 ORIGIN

Query Match 46.58; Score 102.2; DB 11; Length 972;
 Best Local Similarity 69.3%; Pred. No. 4.2e-13;
 Matches 156; Conservative 0; Mismatches 63; Indels 6; Gaps 1;

1 CTATAGTTGATTCCTGTCAGACAAACACACAAAGAGTCACAAAGCAAAATC 60
 83 CTGCGTTGATTTTGCACAGAAATATACCAATGACGAGCCCAACCTACAAAACA 142
 61 AAAATGCAATCTCCGAGACACATGAGTATCGTATTCGAGCATTTAAGAAA 120
 143 AAATGTCGAATCGCAGACACGTGGATGTAATCAAGCGCTGCTGACATTCAGAAAT 202
 121 CCATCCCTATG-----ACAACACAGCCACCAATCCACCTTCATTCGTCAGATTG 174
 203 CCCCTCCATGAGAGACAGAGAGCGCTCCCTCCACCTACATTTTCATTCGTCAGATCCA 262
 175 GACAAAGATTTGCTTTAGTCTTGAACAAATCTGGAAGCATGG 219
 263 GAAGGCGAGTGTGTGCTGCTGCTGATTAATCTGGAAGCATGG 307
 Db
 RESULT 13
 A2855898 744 bp DNA GSS 21-FEB-2001
 LOCUS 2M0160E12F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
 DEFINITION clone UUCG2M0160E12 F, DNA sequence.
 ACCESSION A2855898
 VERSION A2855898.1 GI:13046520
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 744)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10Kb
 plasmid inserts

JOURNAL

Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunne@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0160 row: E column: 12
 Seq primer: CATTGTAAACGACGCGCAGT
 Class: plasmid ends
 High quality sequence stop: 744.
 Location/Qualifiers

FEATURES

source
 1. 744
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_image="UUCG2M0160E12"
 /clone_lib="Mouse 10kb plasmid UUCG1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/nares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pPM2 (p114732114/9b/AF129072.1), a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 239 a 145 c 153 g 207 t
 ORIGIN

Query Match 45.88; Score 100.8; DB 13; Length 744;
 Best Local Similarity 69.4%; Pred. No. 9e-13;
 Matches 154; Conservative 0; Mismatches 62; Indels 6; Gaps 1;

4 TAGTTGATTTCTGACGAACAAACCAACAAAGAGTCACAAAGCAAAATCAA 63
 358 TGGTTGATTTTGCAGAAATATACCAATGACGAGGCCCAACCTACAAAACA 417
 64 AATGCAATCTCCGAGACATGAGAGTATCCCTGATTCGAGACTTTAAGAAAACA 123
 418 TGTGCAATCGAGAGACACGATGATTAATCAAGGCTGTGCTGATTCAGATTC 477
 124 CTCCTATG-----ACAACACAGCCACCAATCCACCTTCTATTCGTCAGATTGAC 177
 478 CTCCTATGAGAGACAGAAACAGAACGCTCCCTCCATCAATTTCAATTCCTCAAGTCAGAA 537
 178 AAAGATTTGCTGTTAGTCTTGAACAAATCTGGAAGCATGG 219
 538 GCGAGTGTGTGCTGCTGCTGATTAATCTGGAAGCATGG 579
 Db
 RESULT 14
 BE656570 402 bp mRNA EST 06-SEP-2000
 LOCUS BE656570
 DEFINITION UT-M-BH0-akf-d-01-0-UI r1 NIH-BMAP_M.S1 Mus musculus cDNA clone
 UT-M-BH0-akf-d-01-0-UI 5', mRNA sequence.
 ACCESSION BE656570


```

VERSION      BE656570.1  GI:3982483
KEYWORDS
SOURCE       house mouse.
ORGANISM     Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 402)
AUTHORS      Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE        Normalization and subtraction: two approaches to facilitate gene
              discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
COMMENT      Contact: Chih, H
              National Institute of Mental Health
              6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
              20892-9643, USA
              Tel: 301 443 1706
              Fax: 301 443 9890
              Email: MEST@mail.nih.gov
              cDNA Library Preparation: M.B. Soares Lab Clone distribution:
              Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
              should be noted that Bento Soares is generating a small number of
              additional specialized non-redundant arrays of BMAP cDNAs whose
              availability will be considered under appropriate and limited
              collaborative arrangements
              Seq primer: M13 Reverse.
FEATURES
  source
    1..402
    /organism="Mus musculus"
    /strain="C57BL/6J"
    /db_xref="taxon:10090"
    /clone="UI-M-BH0-akt-d-01-0-U"
    /clone_1id="NIH_BMAP_M.S1"
    /dev_stage="27-32 days"
    /lab_host="DHI0B (Life Technologies)"
    /note="Vector: pRT3D-Pac (Pharmacia) with a modified
    polylinker. Site_1: Not I; Site_2: Eco RI; The
    NIH_BMAP_M.S1 library is a subtracted library derived from
    a mixture of normalized libraries from ten regions of the
    mouse brain (cerebellum, brain stems, olfactory bulbs,
    hypothalamus, cortex, amygdala, basal ganglia, pineal
    gland, striatum, hippocampus). The driver used for
    subtraction consisted of a pool of 20,000 cDNA clones
    obtained from non-normalized and normalized libraries of
    these ten regions of the mouse brain."
BASE COUNT   136 a          83 g          87 t
ORIGIN
Query Match      45.0%; Score 99; DB 10; Length 402;
Best Local Similarity 68.4%; Pred. No. 2.6e-12;
Matches 154; Conservative 0; Mismatches 65; Indels 6; Gaps 1
  1 CTAAGTTGAAATTCGTCAACAAACCAACAAGACCTCCAAACCAAAATC 60
    1 | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CTGGGTGAAATTTGGCACAGAAAAACCAAAATGCAAGACCCCAACCTACAAACA 60
  61 AAAATGCAATCTCCGAGACACATGGAGAGTATCCGTGATTCGTGAGACTTAAAGAAA 120
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 AAAATGCAATCTCCGAGAGCAGCTGGAGTATTAACAAGCGTCTGCTGACTTTCAGAAATG 120
  121 CCACCTCTATG-----ACAACACAGCCACCAAAATCCACCTTCATTCATTCGTCGACATTTG 174
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 CCCCTCCCATGAGAGACAGAAAGCCCTCCCTCCACCTACATTTTCACATGCTCAAGTCA 180
  175 GACAAAGAAATGTGTGTTAGTCTTGACAAATCTGGAAGCATGG 219
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 GAAAGCGAGTGTGTGCTGTGATTAATTCGGAAGCATGG 225

```

```

DEFINITION      u55f02.y1 NCI CGAP Mam6 Mus musculus cDNA clone IMAGE:3672987 5'
SEQUENCE        similar to TR:088860 088860 CHLONIDE CHANNEL CACG. [I] ;, mRNA
VERSION         BF320704
AUTHORS         EST#
SOURCE          house-mouse.
ORGANISM        Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE       1 (bases 1 to 640)
AUTHORS         NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE           National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                Tumor Gene Index
JOURNAL         Unpublished (1997)
COMMENT         Other-ESTs: u55f02.x1
                Contact: Robert Strausberg, Ph.D.
                Email: cgaps-remail.nih.gov
                Tissue Procurement: Jeffrey Green M.D.
                CDNA Library Preparation: Life Technologies, Inc.
                CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
                DNA Sequencing by: Washington University Genome Sequencing Center
                clome distribution: NCI-CGAP clone distribution information can be
                found through the I.M.A.G.E. Consortium/LNLN at:
                image.lnl.gov/image/html/tresources.shtml

FEATURES             location/Qualifiers
     source           1..640
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3672987"
                     /clone_lib="NCI-CGAP_Mam6"
                     /sex="Female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI. Cloned unidirectionally. Primer: oligo dt.
Library constructed by Life technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT      203 a      154 c      138 g      142 t      3 others
ORIGIN
1 CTATGTGGATTCTGTCAGACAACAAACCACCAACAAAGAAGCTCCAAACGAATAATC 60
||| ||||||| ||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||
Db 177 CTGCGTGTGAATTTTGCCACAGAAAAAACCAACAATGCAGAAAGCCCCAACCACAAAACA 236
Matches 154; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

Query Match      45.0%; Score 99; DB 11; Length 640;
Best Local Similarity 68.4%; Pred. No. 2.3e-12;

QY 11 CTATGTGGATTCTGTCAGACAACAAACCACCAACAAAGAAGCTCCAAACGAATAATC 60
||| ||||||| ||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||
Db 177 CTGCGTGTGAATTTTGCCACAGAAAAAACCAACAATGCAGAAAGCCCCAACCACAAAACA 236
Matches 154; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

QY 61 AAAATGCAATCTCCGAAGCACATGGGAAGTAGTCCGATTCCTGAGAGACTTTAAGAAA 120
||| ||||||| ||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||
Db 237 AAATGTGAATTCGACGAGACAGTGGATGTAATCAAGACGCTGCTGACATTTCAGAAATG 296
||| ||||||| ||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||
QY 121 CCATCCCATG-----ACAACACAGCCACCAACAATCCACCTTCTCATTTGCTGCAGATTG 174
||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 CCCCTCCCATGAGAGAGACAGAAAGGCCCTCTCCACACTACATTTTCACCTGCTCAAGTCCA 356
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 GACAAAGAAATGTGTGTTAGCTTTCCTTGACCAATCGGAAGAGATGG 219
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 357 GAAGCGAGTGTGTGCTTGCTGCTGATTAATCTGGGAAGCATGG 401

```


This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:53:01 ; Search time 521.98 Seconds
(without alignments)
95,454 Million cell updates/sec

Title: US-09-049-696-5

Perfect score: 220
Sequence: 1 CTTAGTGAATTCTGTACA.....GACAAATCTGAAGCATGGC 220

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	29.8	13.5	519	4	US-09-068-140A-3
C 2	29.8	13.5	3447	1	US-08-252-995D-3
C 3	29.8	13.5	3447	2	US-08-834-108-3
C 4	29.6	13.5	1117	4	US-09-247-373B-33
C 5	29.6	13.5	2065	3	US-08-335-865J-8
C 6	29.4	13.4	1140	4	US-08-943-731-209
C 7	29.4	13.4	6243	2	US-09-056-075-1
C 8	29.4	13.4	20084	4	US-08-943-731-5
C 9	28.2	12.8	13146	2	US-08-724-334D-3
C 10	28.2	12.8	13146	3	US-09-270-984A-3
C 11	28.2	12.7	7218	1	US-08-232-463-14
C 12	27.6	12.5	989	4	US-09-312-183A-20
C 13	27.6	12.5	1894	5	PCT-US91-08177-10
C 14	27.6	12.5	2121	2	US-08-897-340-6
C 15	27.6	12.5	2121	3	US-09-252-329-6
C 16	27.6	12.5	7323	5	PCT-US91-08177-1
C 17	27.4	12.5	4711	5	PCT-US95-13658-1
C 18	27.4	12.5	837	1	US-08-832-883-56
C 19	27.4	12.5	837	2	US-08-832-887-56
C 20	27.4	12.5	1910	1	US-08-371-930-2
C 21	27.4	12.5	1910	5	PCT-US94-01712-2
C 22	27.4	12.5	2296	4	US-08-496-841C-137
C 23	27.2	12.4	90050	4	US-09-245-041-5
C 24	27.2	12.4	1798	2	US-08-557-128-12
C 25	26.8	12.2	19557	5	PCT-US92-06300-1
C 26	26.6	12.1	321	2	US-08-520-678A-23
C 27	26.6	12.1	321	4	US-08-897-126-23

C 28	26.6	12.1	356	2	US-08-520-678A-22	Sequence 22, Appl
C 29	26.6	12.1	356	4	US-08-897-126-22	Sequence 22, Appl
C 30	26.6	12.1	1921	2	US-08-557-128-11	Sequence 11, Appl
C 31	26.6	12.1	7859	1	US-07-854-596B-4	Sequence 4, Appl
C 32	26.6	12.1	7859	2	US-08-450-905B-15	Sequence 15, Appl
C 33	26.6	12.1	7859	3	US-07-982-759F-15	Sequence 15, Appl
C 34	26.6	12.1	9646	3	US-08-811-566-1	Sequence 5, Appl
C 35	26.6	12.1	12980	3	US-08-811-566-5	Sequence 5, Appl
C 36	26.4	12.0	2445	6	5215909-9	Patent No. 5215909
C 37	26.4	12.0	246240	2	US-08-724-394A-20	Sequence 20, Appl
C 38	26.4	12.0	246240	2	US-08-724-394A-21	Sequence 21, Appl
C 39	26.4	12.0	246240	2	US-08-724-394A-22	Sequence 22, Appl
C 40	26.2	11.9	260	2	US-08-520-678A-29	Sequence 29, Appl
C 41	26.2	11.9	260	4	US-08-897-126-29	Sequence 29, Appl
C 42	26.2	11.9	317	4	US-09-018-584A-4	Sequence 4, Appl
C 43	26.2	11.9	439	1	US-08-621-502A-6	Sequence 6, Appl
C 44	26.2	11.9	3223	1	US-07-980-528-1	Sequence 1, Appl
C 45	26.2	11.9	51259	3	US-08-781-891-209	Sequence 209, App

ALIGNMENTS

RESULT 1
US-09-068-140A-3/c

Sequence 3, Application US/09068140A

Patent No. 6281409

GENERAL INFORMATION:

APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor

APPLICANT: and Rex Michael Brennan

TITLE OF INVENTION: Blackcurrant Promoters and Genes

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/068,140A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP96/04807

FILING DATE: No. 6281409ember 4, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Dinnert, Data L.

REGISTRATION NUMBER: 33,680

REFERENCE/DOCKET NUMBER: C70237

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5017

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 519 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: CDNA

HYPOTHEICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Ribes nigrum

STRAIN: Ben Alder

US-09-068-140A-3

Query Match 13.5%; Score 29.8; DB 4; Length 519;
Best Local Similarity 50.3%; Pred. No. 1.2;
Matches 73; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
OY 2 TATAGTGAATTTGTGACAGACAAACCAACAAAGAGCTCCAAACAGCAATCA 61
DB 518 TTTTCTTTTCTTTTTCATACATCTACTAGTCTATTATTTTACCACAGAAACACA 459
OY 62 AAATGCAATCTCGAGACGACATGGAGATCCGTGATTCTGAGGACTTTAGAAAC 121
DB 458 AACATGACATTACAGAGACAGCGCAAAACAGCAAAACAGCACTTAGTGCA 399
OY 122 CACTCTATGACACAGACACCA 146
DB 398 TATTCTTTCCACCAACAGACACA 374

RESULT 2
US-08-252-995D-3/C
Sequence 3, Application US/08252995D
Patent No. 5650501
GENERAL INFORMATION:
APPLICANT: Dennis, James W
APPLICANT: Heffernan, Mike
APPLICANT: Fode, Carol
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,995D
FILING DATE: 02-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kurdydyk, Linda M
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Mus musculus
DEVELOPMENTAL STAGE: Lymphoid cDNA Library
IMMEDIATE SOURCE:
LIBRARY: Murine Lymphoid
CLONE: WGA-resistant chop clones
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..205
NAME/KEY: CDS
LOCATION: 206..2980
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 2981..3447

US-08-252-995D-3
Query Match 13.5%; Score 29.8; DB 1; Length 3447;
Best Local Similarity 52.9%; Pred. No. 2.5;
Matches 64; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
OY 16 GTACAGAACAAACCAACAAAGAGCTCCAAACAGCAAAATCAAAATGCAATCTCC 75
DB 3271 GTAAAGAAATATACAAAACAAACAGCCATACACTTAGTTCATACAAACATAGGA 3212
OY 76 GAAGCAGATGAGGAGATGCTGATCTGAGGACTTTAGAAACCACTCTATGACA 135
DB 3211 GCACAAACTTGAAGATATACATCTTGAGCAAAATTAATGATGCTTAGCTTAGACA 3152
OY 136 C 136
DB 3151 C 3151

RESULT 3
US-08-834-108-3/C
Sequence 3, Application US/08834108
Patent No. 5976893
GENERAL INFORMATION:
APPLICANT: Dennis, James W
APPLICANT: Heffernan, Mike
APPLICANT: Fode, Carol
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,108
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kurdydyk, Linda M
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Mus musculus
DEVELOPMENTAL STAGE: Lymphoid cDNA Library
IMMEDIATE SOURCE:
LIBRARY: Murine Lymphoid
CLONE: WGA-resistant chop clones
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..205
NAME/KEY: CDS
LOCATION: 206..2980
FEATURE:

NAME/KEY: 3'UTR
LOCATION: 2981...3447
US-08-834-108-3

Query Match 13.5%; Score 29.8; DB 2; Length 3447;
Best Local Similarity 52.9%; Pred. No. 2.5;
Matches 64; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 16 GTACACACACAAACCAACAAAGACTCCAAACAAACAAATCAAAATGCAATCTCC 75
DB 3271 GTAAAGAAAATAACAAAACAAACCAATACAACTAGTTCATACAAACATAGGA 3212
QY 76 GAAGCAGATGGAGATGATCCGTGATTCGAGACTTTAAGAAACCACTCTATGACAA 135
DB 3211 GCACAAACCTTGAGAGATGATGCTTGAGCAAAATTAAGTCTAGCCTTATGACCA 3152
QY 136 C 136
DB 3151 C 3151

RESULT 4

US-09-247-373B-33
Sequence 33, Application US/09247373B
Patent No. 6168954
GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
APPLICANT: O'KEEFE, DANIEL
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
FILE REFERENCE: CL-1108-A
CURRENT APPLICATION NUMBER: US/09/247,373B
CURRENT FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 08/924,747
PRIOR FILING DATE: 1997-09-05
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Microsoft Office 97
SEQ ID NO 33
LENGTH: 1117
TYPE: DNA
ORGANISM: SOYBEAN
FEATURE:
NAME/KEY: unsure
LOCATION: (1101)
OTHER INFORMATION: M-A OR C
NAME/KEY: unsure
LOCATION: (1104)
OTHER INFORMATION: M-A OR C
NAME/KEY: unsure
LOCATION: (1116)
OTHER INFORMATION: M-G or A or T or C
US-09-247-373B-33

Query Match 13.5%; Score 29.6; DB 4; Length 1117;
Best Local Similarity 64.7%; Pred. No. 1.9;
Matches 44; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 3 ATAGTTGATTCGTACAGACAAACCAACAAAGAGCTCCAAACAAAGCAAAATCA 62
DB 982 ataattcaattcttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1041
QY 63 AATGCAA 70
DB 1042 aaaaaaa 1049

RESULT 5

US-08-335-865J-8
Sequence 8, Application US/08335865J
Patent No. 6107472
GENERAL INFORMATION:
APPLICANT: Stacker, Steven A.; Hovens, Christopher M.,

APPLICANT: Wilks, Andrew F.
TITLE OF INVENTION: RECEPTOR-TYPE TYROSINE KINASE-LIKE MOLECULES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 666 Fifth Ave
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720 KB storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: ASCII/wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/335,865J
FILING DATE: 19-January-1995

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU93/00210
FILING DATE: 10-May-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PL2358
FILING DATE: 11-May-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6107472man D.
REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD-5277
TELEPHONE: (212) 318-3100
TELEFAX: (212) 752-5958

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2065
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-335-865J-8

Query Match 13.5%; Score 29.6; DB 3; Length 2065;
Best Local Similarity 46.6%; Pred. No. 2.4;
Matches 95; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 129 GATTTCTGTACAGACAAACCAACAAAGAGCTCCAAACAAAGCAAAATCAAAATGC 68
DB 1102 GATTTAGTAGATGAAAGAAAGACCAATTAAGAGCAAAACATTTGTAAACAGTTAA 1161
QY 59 AATCTCCGAAGCAGATGGAGATGATCGTGTATCTGAGGACTTTAAGAAACCACTCCT 128
DB 1152 AGACCAAGCATCTGAGATCGATGAGTGCATGATGCTACCCGAGAGCTCCAGCTTGAGG 1221
QY 129 ATGACACACAGCCACCAATCCACCTTCTCATCTTCCTCAGATTTGAGACAAAGATGTG 188
DB 1252 TCTGACACCAAGAAACCTCTCTTCTATCTATGTGTCATGAGAGAGAAAGCAAGCC 1281
QY 189 TGTTCAGCTCTGACCAATCTGGA 212
DB 1232 CATGTGTATGCTCATACATGA 1305

RESULT 6

US-08-943-731-209/C
Sequence 209, Application US/08943731
Patent No. 6265157
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL


```

COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056.075
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960236.95238
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
POPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3770..4013
OTHER INFORMATION: /note="RP4 origin of DNA transfer (oritL) from
OTHER INFORMATION: plasmid Rp4"
US-09-056-075-1

Query Match 13.4%; Score 29.4; DB 2; Length 6243;
Best Local Similarity 52.9%; Pred. No. 4.2;
Matches 63; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

CY 2 TATAGTTAATTCTGTACAGACAAACACACAAAGAGCTCCAAACAGCAAAATCA 61
D 3313 TATATCAAACTCGTAAAGTAAAGAAATAAATTTCTTAAAAATTAAGTTAAATA 3254
DB 3253 AGCCGGAATTTCACTAGTTTAAAAAACATATAATATAAAGATATTATTATA 3195

RESULT: 8
US-08-943-731-5/C
Sequence 5, Application US/08943731
Patent No. 6265157
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISSA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORRKO, JARMO
APPLICANT: ALA-KORRKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
STREET: FLR.
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:

```


Query Match	12.8%	Score 28.2	DB 3	Length 13146
Best Local Similarity	59.3%	Pred. No. 14		
Matches 48	Conservative 0	Mismatches 33	Indels 0	Gaps 0

OY 9 GAATTCGTACAGACAAACCAACAAAGAGCTCCAAAGCAAAATCAAAATGC 68
DB 12782 GAAACGCAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGCA 12723
OY 69 AATCTCCGAGACATGGGAA 89
DB 12722 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 12702

RESULT 11
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 12.7%; Score 28; DB 1; Length 7218;
Best Local Similarity 11.0%; Pred. No. 12;
Matches 13; Conservative 65; Mismatches 40; Indels 0; Gaps 0;

OY 3 ATAGTTGAATTCGTACAGACAAACCAACAAAGAGCTCCAAAGCAAAATCAAA 62
DB 1451 ATAGAGAAATTTGTACACACACACACACACACACACACACACACACACAC 1392
OY 63 AATGCAATCTCCGAGACATGGGAA 120
DB 1391 RRR 1334

RESULT 12

US-09-312-183A-20
Sequence 20, Application US/09312183A
Patent No. 6303766
GENERAL INFORMATION:
APPLICANT: GRABAU, ELIZABETH A.
APPLICANT: HEGEMAN, CARLA E.
TITLE OF INVENTION: SOYBEAN PHYTASE AND NUCLEIC ACID ENCODING THE SAME
FILE REFERENCE: 6617-13
CURRENT APPLICATION NUMBER: US/09/312.183A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 20
LENGTH: 989
TYPE: DNA
ORGANISM: Glycine max
US-09-312-183A-20

Query Match 12.5%; Score 27.6; DB 4; Length 989;
Best Local Similarity 52.6%; Pred. No. 7.7;
Matches 60; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

OY 11 ATCTGTACAGACAAACCAACAAAGAGCTCCAAAGCAAAATCAAAATGCA 70
DB 393 atagagtaagaaataaataaactaaagagtgagaaataaataaataaagcga 452
OY 71 TCTCCGAGACATGGGAGAGTCCGTGATCTCGAGACTTTAGAAAAACAC 124
DB 453 tgttcaatctactcttgacggagaaataatgacttaagaactgaataagaagac 506

RESULT 13
PCT-US91-08177-10.

Sequence 10, Application PC/TUS9108177
GENERAL INFORMATION:
APPLICANT: Samal, Siba K
TITLE OF INVENTION: Bovine Respiratory Syncytial Virus Genes
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08177
FILING DATE: 19911104
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/608,937
FILING DATE: 05-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Higbet, David W
REGISTRATION NUMBER: 30,265
REFERENCE/DOCKET NUMBER: 20509-96711
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4854
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1894 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO

ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bovine respiratory syncytial virus
STRAIN: A 51908
FEATURE:
NAME/KEY: CDS
LOCATION: 14..1729
OTHER INFORMATION: /label= F gene
PCT-US91-0817-10

Query Match 12.5%; Score 27.6; DB 5; Length 1894;
Best Local Similarity 58.5%; Pred. No. 10;
Matches 48; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 7 TTGAATTCGTACAGAACCAACCAAGACCTCCCAACCAAGCAAAATCAAAAT 66
DB 232 TTCMAAGTGAATTAATTAAGCAAGACTAGAAAGATCAACATGCACTAGCGAAT 291
QY 67 GCAATCTCCGAAGCAGCATGGGA 88
DB 292 GCAATCATTATGCAAAATGAA 313

RESULT 14

US-08-897-340-6/C
Sequence 6, Application US/08897340
Patent No. 5955306

GENERAL INFORMATION:

APPLICANT: Gimeno, Carlos J. and Errada, Patrick, R.
TITLE OF INVENTION: Weight Control Pathway Genes and Uses
TITLE OF INVENTION: Therefor
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,340
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/715,032
FILING DATE: 17-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Silverl, Jean M.
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: MNI-005CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2121 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-897-340-6

Query Match 12.5%; Score 27.6; DB 2; Length 2121;
Best Local Similarity 51.6%; Pred. No. 10;
Matches 63; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 10 AATCTGTACAGAACCAACCAAGACCTCCCAACCAAGCAAAATCAAAATGCA 69
DB 1660 AATCTGTACATACACTGCAAACTCTAGCAAAAGATATATGTATATATACATGCA 1601
QY 70 ATCTCGAAGCAGATGGGAAGTATCTCGATTTCTAGAGACTTTAGAAACACCTCTTA 129
DB 1680 AAGTGTATGTACATATCCACATACATTCGCAACTTTGGATTTTAAAAACTAATTGAA 1541
QY 190 TG 131
DB 1530 TG 1539

RESULT 15

US-09-252-329-6/C
Sequence 6, Application US/09252329
Patent No. 6147192

GENERAL INFORMATION:

APPLICANT: Gimeno, Carlos J. and Errada, Patrick, R.
TITLE OF INVENTION: Weight Control Pathway Genes and Uses
TITLE OF INVENTION: Therefor
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/252,329
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/897,340
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Silverl, Jean M.
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: MNI-005CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2121 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-252-329-6

Query Match 12.5%; Score 27.6; DB 3; Length 2121;
Best Local Similarity 51.6%; Pred. No. 10;
Matches 63; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 10 AATCTGTACAGAACCAACCAAGACCTCCCAACCAAGCAAAATCAAAATGCA 69
DB 1660 AATCTGTACATACACTGCAAACTCTAGCAAAAGATATATGTATATATACATGCA 1601
QY 70 ATCTCGAAGCAGATGGGAAGTATCTCGATTTCTAGAGACTTTAGAAACACCTCTTA 129
DB 1690 AAGTGTATGTACATATCCACATACATTCGCAACTTTGGATTTTAAAAACTAATTGAA 1541
QY 130 TG 131
DB 1530 TG 1539

Thu Apr 4 09:27:46 2002

Search completed: April 3, 2002, 20:53:10
Job time: 41748 sec

us-09-049-696-5.rml

Page 8

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:18:10 ; Search time 1321.64 Seconds
(without alignments)
147.251 Million cell updates/sec

Title: US-09-049-696-13

Sequence: 1 GTGGCGGCTCTGGAGAGT.....AATGCTCCATCACTGATCT 227

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

N.Ceneseq_1101.*
1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	227	100.0	878	18	AA45884 Human colon specif
2	227	100.0	878	19	AAV16672 POLYNUCLEOTIDE seq
3	227	100.0	2742	22	AAH46102 Human ICACC coding
4	227	100.0	2745	20	AA81927 Human ICACC-1 nucl
5	227	100.0	2825	22	AAH46124 Human ICACC-1 nucl
6	227	100.0	2854	22	AAH34879 Human colon cancer
7	227	100.0	2854	22	AAH34879 Human colon cancer
8	227	100.0	2867	22	AAH33285 Human colon cancer
9	227	100.0	3109	22	AAH35019 Human colon cancer
10	227	100.0	3111	20	AA209840 Human membrane spa
11	227	100.0	3311	22	AA129502 C902P determined c

12	104.8	46.2	2739	22	AAH46101	Mouse Gob-5 coding
13	104.8	46.2	2843	22	AAH46120	Mouse Gob-5 cDNA,
14	104.8	46.2	2931	20	AA81925	Murine ICACC-1 nuc
15	94.8	41.8	2616	21	AAH64335	Clone 2516888 of a
16	94.8	41.8	3265	21	AAH65095	Membrane-bound pro
17	94.8	41.8	3265	22	AA82092	Human PRO1124 cDNA
18	94.8	41.8	3265	22	AA82092	Human PRO1124 (UNG
19	89.8	39.6	1802	22	AA82092	Human colon cancer
20	89.8	39.6	1802	22	AAH33192	Human colon cancer
21	88.6	39.0	576	21	AAH16018	Human colon cancer
22	80.6	35.5	611	21	AAH16022	Human colon cancer
23	67.8	29.9	742	21	AAH16028	Human colon cancer
24	66.8	29.4	595	21	AAH16020	Human colon cancer
25	33.2	14.6	486	22	AAH46121	Mouse Gob-5 cDNA h
26	33.2	14.6	2407	22	AAH18511	Human cDNA sequenc
27	33.2	14.6	2454	22	AAH82881	Human tumor-associ
28	33.2	14.6	2784	20	AAH24658	Human lung tumor a
29	33.2	14.6	2784	20	AAH24658	Human lung tumor a
30	33.2	14.6	3156	21	AAH51625	Human lung cancer-
31	33.2	14.6	3951	21	AAH24653	Human lung tumor a
32	33.2	14.6	3951	21	AAH24653	Human lung cancer-
33	33.2	14.6	8031	21	AAH65950	Human lung cancer-
34	33.2	14.5	1158	21	AAH15719	Human prostate can
35	33.2	14.5	8034	20	AAH88525	Human prostate can
36	31.4	13.8	1096	22	AAH20247	Candida albicans a
37	31.4	13.8	1106	22	AAH24224	DNA encoding molec
38	31.4	13.8	6760	21	AAH28293	Human oxidoreducta
39	31.2	13.7	9960	21	AAH250905	Sorbum resistance
40	30.2	13.3	2720	22	AAH15935	Human TBC-1 partia
41	30.2	13.2	129	22	AAH10009	Human cDNA sequenc
42	29.8	13.1	469	22	AAH105038	Probe #5029 used t
43	29.6	13.0	14460	21	AAH293815	Olfactory receptor
44	29.4	13.0	1146	20	AAH37374	Human secreted pro
45	29.4	13.0	4931	21	AAH96639	Virulence gene #46

ALIGNMENTS

RESULT 1	AA45884 standard; cDNA; 878 BP.
1	AA45884
2	AA45884
3	AA45884
4	AA45884
5	AA45884
6	AA45884
7	AA45884
8	AA45884
9	AA45884
10	AA45884
11	AA45884
12	AA45884
13	AA45884
14	AA45884
15	AA45884
16	AA45884
17	AA45884
18	AA45884
19	AA45884
20	AA45884
21	AA45884
22	AA45884
23	AA45884
24	AA45884
25	AA45884
26	AA45884
27	AA45884
28	AA45884
29	AA45884
30	AA45884
31	AA45884
32	AA45884
33	AA45884
34	AA45884
35	AA45884
36	AA45884
37	AA45884
38	AA45884
39	AA45884
40	AA45884
41	AA45884
42	AA45884
43	AA45884
44	AA45884
45	AA45884
46	AA45884
47	AA45884
48	AA45884
49	AA45884
50	AA45884
51	AA45884
52	AA45884
53	AA45884
54	AA45884
55	AA45884
56	AA45884
57	AA45884
58	AA45884
59	AA45884
60	AA45884
61	AA45884
62	AA45884
63	AA45884
64	AA45884
65	AA45884
66	AA45884
67	AA45884
68	AA45884
69	AA45884
70	AA45884
71	AA45884
72	AA45884
73	AA45884
74	AA45884
75	AA45884
76	AA45884
77	AA45884
78	AA45884
79	AA45884
80	AA45884
81	AA45884
82	AA45884
83	AA45884
84	AA45884
85	AA45884
86	AA45884
87	AA45884
88	AA45884
89	AA45884
90	AA45884
91	AA45884
92	AA45884
93	AA45884
94	AA45884
95	AA45884
96	AA45884
97	AA45884
98	AA45884
99	AA45884
100	AA45884
101	AA45884
102	AA45884
103	AA45884
104	AA45884
105	AA45884
106	AA45884
107	AA45884
108	AA45884
109	AA45884
110	AA45884
111	AA45884
112	AA45884
113	AA45884
114	AA45884
115	AA45884
116	AA45884
117	AA45884
118	AA45884
119	AA45884
120	AA45884
121	AA45884
122	AA45884
123	AA45884
124	AA45884
125	AA45884
126	AA45884
127	AA45884
128	AA45884
129	AA45884
130	AA45884
131	AA45884
132	AA45884
133	AA45884
134	AA45884
135	AA45884
136	AA45884
137	AA45884
138	AA45884
139	AA45884
140	AA45884
141	AA45884
142	AA45884
143	AA45884
144	AA45884
145	AA45884
146	AA45884
147	AA45884
148	AA45884
149	AA45884
150	AA45884
151	AA45884
152	AA45884
153	AA45884
154	AA45884
155	AA45884
156	AA45884
157	AA45884
158	AA45884
159	AA45884
160	AA45884
161	AA45884
162	AA45884
163	AA45884
164	AA45884
165	AA45884
166	AA45884
167	AA45884
168	AA45884
169	AA45884
170	AA45884
171	AA45884
172	AA45884
173	AA45884
174	AA45884
175	AA45884
176	AA45884
177	AA45884
178	AA45884
179	AA45884
180	AA45884
181	AA45884
182	AA45884
183	AA45884
184	AA45884
185	AA45884
186	AA45884
187	AA45884
188	AA45884
189	AA45884
190	AA45884
191	AA45884
192	AA45884
193	AA45884
194	AA45884
195	AA45884
196	AA45884
197	AA45884
198	AA45884
199	AA45884
200	AA45884
201	AA45884
202	AA45884
203	AA45884
204	AA45884
205	AA45884
206	AA45884
207	AA45884
208	AA45884
209	AA45884
210	AA45884
211	AA45884
212	AA45884
213	AA45884
214	AA45884
215	AA45884
216	AA45884
217	AA45884
218	AA45884
219	AA45884
220	AA45884
221	AA45884
222	AA45884
223	AA45884
224	AA45884
225	AA45884
226	AA45884
227	AA45884
228	AA45884
229	AA45884
230	AA45884
231	AA45884
232	AA45884
233	AA45884
234	AA45884
235	AA45884
236	AA45884
237	AA45884
238	AA45884
239	AA45884
240	AA45884
241	AA45884
242	AA45884
243	AA45884
244	AA45884
245	AA45884
246	AA45884
247	AA45884
248	AA45884
249	AA45884
250	AA45884
251	AA45884
252	AA45884
253	AA45884
254	AA45884
255	AA45884
256	AA45884
257	AA45884
258	AA45884
259	AA45884
260	AA45884
261	AA45884
262	AA45884
263	AA45884
264	AA45884
265	AA45884
266	AA45884
267	AA45884
268	AA45884
269	AA45884
270	AA45884
271	AA45884
272	AA45884
273	AA45884
274	AA45884
275	AA45884
276	AA45884
277	AA45884
278	AA45884
279	AA45884
280	AA45884
281	AA45884
282	AA45884
283	AA45884
284	AA45884
285	AA45884
286	AA45884
287	AA45884
288	AA45884
289	AA45884
290	AA45884
291	AA45884
292	AA45884
293	AA45884
294	AA45884
295	AA45884
296	AA45884
297	AA45884
298	AA45884
299	AA45884
300	AA45884
301	AA45884
302	AA45884
303	AA45884
304	AA45884
305	AA45884
306	AA45884

XX		(TAKE) TAKEDA CHEM IND LTD.	
PA			
XX		Nakanishi A, Morita S;	
PI			
DR	WI:	2001-355935/37.	
XX	P-PSDB:	AAB73716.	
PR		New antisense nucleotide, useful for treatment and prevention of	
PT		bronchial asthma and chronic obstructive pulmonary disease	
XX			
PS	Claim 3; Page 82-84; 104pp:	Japanese.	
XX			
CC	The invention relates to an antisense nucleotide targeted to the mouse		
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,		
CC	the CLCA1 gene (coding sequence shown in AAH46102). The invention also		
CC	relates to an antibody specific for the Gob-5 protein, medical and		
CC	diagnostic compositions containing the antisense nucleotide or the		
CC	antibody, and methods and kits for screening for compounds which inhibit		
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.		
CC	The antisense oligonucleotides and antibody are therefore useful for the		
CC	treatment and prevention of bronchial asthma and chronic obstructive		
CC	pulmonary disease. The present sequence represents the human CLCA1		
CC	gene coding sequence.		
XX			
SQ	Sequence 2742 BP: 833 A; 616 C; 623 G; 670 T; 0 other:		
	Query Match	100.0%; Score 227; DB 22; Length 2742;	
	Best Local Similarity	100.0%; Pred. No. 1e-67;	
	Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 GTGCGGGCTCTGGGAGGAAGTTAACGCACGACGAGCAGATACCCACGACACTGA 60		
DB	2023 gTgCGgggctcTgTggggagTtaacgcagcgacgaaggagtataccacgaagtgTga 2082		
OY	61 GCACGTGTCATRCCTGGCTGGCATGTAGAATGATGAATCAATGGAATCCACCAGACT 120		
DB	2083 gcactgtacatacccttgcgtgagttagaatgatgaatatcaatggaatccaccaagacct 2142		
OY	121 GAATTAAATAAGATGATGATGTTCAACAACAAGCAAGTGTTTTAGCAGCAATCCTCGGGA 180		
DB	2143 gaattaaataagaTgaTgattcaacaacaagaagTgTttcaagaagaacatcccgcggga 2202		
OY	181 GGCTCATTTTGGGCTTCGTGATGTGCCAATGCTCCCATACCTGATCT 227		
DB	2203 ggctcatcttgTgcttcgtatgtgccaatgtgtcccatcctcgtacct 2249		
RESULT	4		
ID	AAF81927		
XX	AAF81927 standard; cDNA: 2745 BP.		
XX	AAF81927:		
DT	13-JUN-2001 (first entry)		
XX			
DE	Human ICACC-1 nucleotide sequence.		
XX			
KW	ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;		
KW	Interleukin 9 induced calcium activated chloride channel; IL-9;		
KW	calcium activated chloride channel; anti-allergic; anti-asthmatic;		
KW	anti-inflammatory; immunomodulatory; cystic fibrosis;		
XX	Inflammatory bowel disease; autoimmune disease; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..2745	
FT		/tag=a	
FT		/product="ICACC-1"	
FT		/note="IL-9 induced calcium activated chloride channel"	
XX			

PN		MO9944620-A1	
PD		10-SEP-1999.	
XX			
XX			
FE	U3-MAR-1999;	99WO-US04703.	
XX			
PR	03-MAR-1998;	98US-0076815.	
XX			
PA	(MAGA-) MAGAININ PHARM INC.		
XX	Holtroyd RJ, Levitt RC, Maloy WL, Loushed J, McLane M;		
PI	Nicolaides NC, Zhou Y, Dong Q;		
DR	WPE; 1999-550979/46.		
P	F-PSDB; AAB74824.		
HT	New nucleic acid encoding calcium activated chloride channel, used to		
FT	identify, e.g. specific modulators for treating atopic allergy -		
XX			
PS	Clim 1; Fig 4B; 75pp; English.		
XX			
CC	The present sequence encodes the human interleukin 9 (IL-9) induced		
CC	calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins		
CC	have anti-allergic, anti-asthmatic, anti-inflammatory and		
CC	immunomodulatory activities. Compounds (A) that downregulate ICACC are		
CC	used to alleviate asthma (or more generally atopic allergy), while those		
CC	(B) that upregulate ICACC are used to treat cystic fibrosis (CF) and		
CC	inflammatory bowel disease (IBD) (or other autoimmune diseases).		
CE	Measuring levels of ICACC-1 can be used in the diagnosis of asthma		
CC	(increased levels) or IBD (reduced levels), also for monitoring		
CC	treatment of these conditions. The ICACC proteins can be used:		
CC	(1) to raise specific antibodies (Ab), useful: (a) as immunoassay		
CC	reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as		
CC	(A) to disrupt bonding between ICACC and its ligands; and (iii) to		
CC	identify modulators and binding partners. ICACC polynucleotides can be		
CC	used to generate transgenic animals or recombinant cells, used to screen		
CC	for antagonists, also as a source of therapeutic antisense agents or		
CC	diagnostic probes (for quantifying mRNA expression, e.g. for		
CC	identification of modulators).		
XX			
SQ	Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other:		
J			
Query_Match	100.0%; Score 227; DB 20; Length 2745;		
Best Local Similarity	100.0%; Pred. No. 1e-67;		
Matches	227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 GTGGGCGCTCTGTGGAGGACTTACGCACGCCAACAGCAGAGTAGTATACCCACGACGACTGA 60		
FD	2093 gTgcgggctcTtggaaggatlaacgaagccagaagcgagtagatacccacgaagtgga 2082		
OY	1 GCACGTCATCATCCGCGCGGATGAGATGATGTAATCAATGCAATGCATTACCAAGACT 120		
DB	2093 gaacttcatcaaccggtcggtgagtgaagaattgtaaatcaatgaatcaccaagacct 2142		
OY	121 GAATAATTAAGATGATGTTCAACAACAAGCAAGTGTTGTTCAGAGAAATCTCTGGGA 180		
PD	2113 gaattataaagaaagatgattcaacaacaagaagtggtttcaagcagaacatccctcgga 2202		
OY	111 GGCTCATTTGTGGCTTGTGATGTCCCAATGCTCCCAATACCTGATCT 227		
DB	2203 ggctcatcttggtgcttcgatgtgcccaaatgctcccatlaccatcgatct 2249		
RESULT	5		
FAH46124			
ID	FAH46124 standard; cDNA; 2825 BP.		
XX	AAG46124;		
XX			
XX	11 SEP-2001 (first entry)		
XX			
DE	Human CLCA1 cDNA, SEQ ID NO:26.		


```
XX Human CLCA1: goblet cell; mouse Gob-5 orthologue; drug screening;
KW expression inhibition; antisense therapy; gene therapy;
KW chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
KW ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 23..2767
XX /tag= a
XX /product= "Human CLCA1"
XX /transl_except= (pos:476..478, aa:lys)
XX WO200138530-A1.
XX 31-MAY-2001.
XX 22-NOV-2000: 2000WO-JP08232.
XX 24-NOV-1999: 99JP-0333479.
XX 27-APR-2000: 2000JP-0127589.
XX (TAKE ) TAKEDA CHEM IND LTD.
XX Nakanishi A, Morita S;
XX WPI; 2001-355935/37.
XX P-PSDB; AAB3716.
XX New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease -
XX Example 5; Page 92-94; 104pp; Japanese.
XX The invention relates to an antisense nucleotide targeted to the mouse
XX Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX relates to an antibody specific for the Gob-5 protein, medical and
XX diagnostic compositions containing the antisense nucleotide or the
XX antibody, and methods and kits for screening for compounds which inhibit
XX the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX The antisense oligonucleotides and antibody are therefore useful for the
XX treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents human CLCA1 cDNA.
XX Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;
SQ
Query Match 100.0%; Score 227; DB 22; Length 2825;
Best Local Similarity 100.0%; Pred. No. 1e-67;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GTGGGGGCTCTGGAGAGTTAAGCAGCCAGAGAGATGATACCCGAGAGATGGA 60
DB 2045 gtgcgggctctggagagtttaagcagccagagagatgatacccgagagtgga 2104
OY 61 GCACGTGATACCTGGCTGGATGAGATGATGAATATCAATGGAATCCACCAAGACT 120
DB 2105 gcaactgtacactggtctggtatgagagatgataacatacctggaatcccaagact 2164
OY 121 GAAATTAATAGATGATGTTCAACACAGCAAGTGTGTTTCAGCAGANACTCTGGGA 180
DB 2165 gaatttaataagatgattgttcaacacagcaagtgtgttccagcagaaacatcctcgga 2224
OY 181 GGCTCATTTGGCTTGTGATGTCCCAATGCTCCCATCTGATCT 227
DB 2225 ggcctcatTTGGCTTGTGATGTCCCAATGCTCCCATCTGATCT 2271
RESULT 6
AAH34879
ID AAH34879 standard; cDNA: 2854 BP.
```

```
XX AAH34879;
XX AC 03-SEP-2001 (first entry)
XX DT Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
XX DE Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX KW colorectal carcinoma; chromosome 1; ss.
XX Homo sapiens;
XX WO200122920-A2.
XX 05-APR-2001;
XX 28-SEP-2000; 2000WO-US26524.
XX 29-SEP-1999; 99US-0157137.
XX 03-NOV-1999; 99US-0163280.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
XX F-PSDB; AAG75474.
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX Claim 1; Page 3462-3463; 9803pp; English.
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing
XX inactive proteins or to supplement the patient's own production of P.
XX Additionally, N may be used to produce the colon cancer-associated P,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAG77789 represent sequences used in the exemplification of the
XX present invention.
XX K.O. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.
XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;
SQ
Query Match 100.0%; Score 227; DB 22; Length 2854;
Best Local Similarity 100.0%; Pred. No. 1e-67;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GTGGGGGCTCTGGAGAGTTAAGCAGCCAGAGAGATGATACCCGAGAGATGGA 60
DB 2057 gtgcgggctctggagagtttaagcagccagagagatgatacccgagagtgga 2116
OY 61 GCACGTGATACCTGGCTGGATGAGATGATGAATATCAATGGAATCCACCAAGACT 120
DB 2117 gcaactgtacactggtctggtatgagagatgataacatacctggaatcccaagact 2176
OY 121 GAAATTAATAGATGATGTTCAACACAGCAAGTGTGTTTCAGCAGANACTCTGGGA 180
DB 2177 gaatttaataagatgattgttcaacacagcaagtgtgttccagcagaaacatcctcgga 2236
OY 181 GGCTCATTTGGCTTGTGATGTCCCAATGCTCCCATCTGATCT 227
ID AAH34879 standard; cDNA: 2854 BP.
```


Db 2237 ggcctcattgtggtctgtatgctcccaaaagctcccaactgattct 2283

RESULT 7

AAF81787 AAF81787 standard; cDNA; 2854 BP.

AC AAF81787;

DT 12-JUN-2001 (first entry)

DE Human secreted protein gene 1 SEQ ID NO:11.

XX Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
XX dermatological; immunosuppressive; antiinflammatory; anti-HIV;
XX immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;
XX ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
XX antialzheimer; antiparkinsonian; antimicrobial; vulnary; gene therapy;
XX immune disorder; hyperproliferative disorder; cardiovascular disease;
XX cancer; angiogenic disorder; neurological disorder; infectious disease;
XX wound healing; regeneration; chemotaxis; chromosome 1; ss.

OS Homo sapiens.

PN WO200112775-A2.

PD 22-FEB-2001.

PF 16-AUG-2000; 2000WO-US22325.

PR 17-AUG-1999; 99US-0149182.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ni J, Florence KA, Fisceella M, Wei P, Baker KP;

PI Birse CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;

DR P-PSDB; AAB74733.

PT Nucleic acids encoding 25 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -

PS Claim 1; Page 441; 485bp; English.

XX AAF81787 to AAF81817 encode the human secreted proteins given in AAB74733
CC to AAB74772. Human secreted proteins can have activities based on the
CC tissues and cells they are expressed in. Example of activities include:
CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;
CC antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;
CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
CC anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and
CC vulnary. Human secreted proteins can be used in gene therapy and
CC vaccine. Human secreted protein nucleotide sequences (NMI) and proteins
CC (PEP) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. For example, NMI
CC and PEP may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of proteins by expressing inactive proteins or
CC to supplement the patient's own production of polypeptides. Disorders that
CC may be prevented, diagnosed and/or treated include immune disorders,
CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
CC angiogenic disorders, neurological disorders, infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to
CC AAF81786 and AAB74732 represent sequences used in the exemplification of
CC the present invention.

SO Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match 100.0%; Score 227; DB 22; Length 2854;

Best Local Similarity 100.0%; Pred. No. 16-67;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

161 GTGGGGCTCTGGGAGAGTTAAGCGACCGACAGGAGAGTGAATCCCGACAGAGTGA 60
162 |
163 |
164 |
165 |
166 |
167 |
168 |
169 |
170 |
171 |
172 |
173 |
174 |
175 |
176 |
177 |
178 |
179 |
180 |
181 |
182 |
183 |
184 |
185 |
186 |
187 |
188 |
189 |
190 |
191 |
192 |
193 |
194 |
195 |
196 |
197 |
198 |
199 |
200 |
201 |
202 |
203 |
204 |
205 |
206 |
207 |
208 |
209 |
210 |
211 |
212 |
213 |
214 |
215 |
216 |
217 |
218 |
219 |
220 |
221 |
222 |
223 |
224 |
225 |
226 |
227 |
228 |
229 |
230 |
231 |
232 |
233 |
234 |
235 |
236 |
237 |
238 |
239 |
240 |
241 |
242 |
243 |
244 |
245 |
246 |
247 |
248 |
249 |
250 |
251 |
252 |
253 |
254 |
255 |
256 |
257 |
258 |
259 |
260 |
261 |
262 |
263 |
264 |
265 |
266 |
267 |
268 |
269 |
270 |
271 |
272 |
273 |
274 |
275 |
276 |
277 |
278 |
279 |
280 |
281 |
282 |
283 |
284 |
285 |
286 |
287 |
288 |
289 |
290 |
291 |
292 |
293 |
294 |
295 |
296 |
297 |
298 |
299 |
300 |
301 |
302 |
303 |
304 |
305 |
306 |
307 |
308 |
309 |
310 |
311 |
312 |
313 |
314 |
315 |
316 |
317 |
318 |
319 |
320 |
321 |
322 |
323 |
324 |
325 |
326 |
327 |
328 |
329 |
330 |
331 |
332 |
333 |
334 |
335 |
336 |
337 |
338 |
339 |
340 |
341 |
342 |
343 |
344 |
345 |
346 |
347 |
348 |
349 |
350 |
351 |
352 |
353 |
354 |
355 |
356 |
357 |
358 |
359 |
360 |
361 |
362 |
363 |
364 |
365 |
366 |
367 |
368 |
369 |
370 |
371 |
372 |
373 |
374 |
375 |
376 |
377 |
378 |
379 |
380 |
381 |
382 |
383 |
384 |
385 |
386 |
387 |
388 |
389 |
390 |
391 |
392 |
393 |
394 |
395 |
396 |
397 |
398 |
399 |
400 |
401 |
402 |
403 |
404 |
405 |
406 |
407 |
408 |
409 |
410 |
411 |
412 |
413 |
414 |
415 |
416 |
417 |
418 |
419 |
420 |
421 |
422 |
423 |
424 |
425 |
426 |
427 |
428 |
429 |
430 |
431 |
432 |
433 |
434 |
435 |
436 |
437 |
438 |
439 |
440 |
441 |
442 |
443 |
444 |
445 |
446 |
447 |
448 |
449 |
450 |
451 |
452 |
453 |
454 |
455 |
456 |
457 |
458 |
459 |
460 |
461 |
462 |
463 |
464 |
465 |
466 |
467 |
468 |
469 |
470 |
471 |
472 |
473 |
474 |
475 |
476 |
477 |
478 |
479 |
480 |
481 |
482 |
483 |
484 |
485 |
486 |
487 |
488 |
489 |
490 |
491 |
492 |
493 |
494 |
495 |
496 |
497 |
498 |
499 |
500 |
501 |
502 |
503 |
504 |
505 |
506 |
507 |
508 |
509 |
510 |
511 |
512 |
513 |
514 |
515 |
516 |
517 |
518 |
519 |
520 |
521 |
522 |
523 |
524 |
525 |
526 |
527 |
528 |
529 |
530 |
531 |
532 |
533 |
534 |
535 |
536 |
537 |
538 |
539 |
540 |
541 |
542 |
543 |
544 |
545 |
546 |
547 |
548 |
549 |
550 |
551 |
552 |
553 |
554 |
555 |
556 |
557 |
558 |
559 |
560 |
561 |
562 |
563 |
564 |
565 |
566 |
567 |
568 |
569 |
570 |
571 |
572 |
573 |
574 |
575 |
576 |
577 |
578 |
579 |
580 |
581 |
582 |
583 |
584 |
585 |
586 |
587 |
588 |
589 |
590 |
591 |
592 |
593 |
594 |
595 |
596 |
597 |
598 |
599 |
600 |
601 |
602 |
603 |
604 |
605 |
606 |
607 |
608 |
609 |
610 |
611 |
612 |
613 |
614 |
615 |
616 |
617 |
618 |
619 |
620 |
621 |
622 |
623 |
624 |
625 |
626 |
627 |
628 |
629 |
630 |
631 |
632 |
633 |
634 |
635 |
636 |
637 |
638 |
639 |
640 |
641 |
642 |
643 |
644 |
645 |
646 |
647 |
648 |
649 |
650 |
651 |
652 |
653 |
654 |
655 |
656 |
657 |
658 |
659 |
660 |
661 |
662 |
663 |
664 |
665 |
666 |
667 |
668 |
669 |
670 |
671 |
672 |
673 |
674 |
675 |
676 |
677 |
678 |
679 |
680 |
681 |
682 |
683 |
684 |
685 |
686 |
687 |
688 |
689 |
690 |
691 |
692 |
693 |
694 |
695 |
696 |
697 |
698 |
699 |
700 |
701 |
702 |
703 |
704 |
705 |
706 |
707 |
708 |
709 |
710 |
711 |
712 |
713 |
714 |
715 |
716 |
717 |
718 |
719 |
720 |
721 |
722 |
723 |
724 |
725 |
726 |
727 |
728 |
729 |
730 |
731 |
732 |
733 |
734 |
735 |
736 |
737 |
738 |
739 |
740 |
741 |
742 |
743 |
744 |
745 |
746 |
747 |
748 |
749 |
750 |
751 |
752 |
753 |
754 |
755 |
756 |
757 |
758 |
759 |
760 |
761 |
762 |
763 |
764 |
765 |
766 |
767 |
768 |
769 |
770 |
771 |
772 |
773 |
774 |
775 |
776 |
777 |
778 |
779 |
780 |
781 |
782 |
783 |
784 |
785 |
786 |
787 |
788 |
789 |
790 |
791 |
792 |
793 |
794 |
795 |
796 |
797 |
798 |
799 |
800 |
801 |
802 |
803 |
804 |
805 |
806 |
807 |
808 |
809 |
810 |
811 |
812 |
813 |
814 |
815 |
816 |
817 |
818 |
819 |
820 |
821 |
822 |
823 |
824 |
825 |
826 |
827 |
828 |
829 |
830 |
831 |
832 |
833 |
834 |
835 |
836 |
837 |
838 |
839 |
840 |
841 |
842 |
843 |
844 |
845 |
846 |
847 |
848 |
849 |
850 |
851 |
852 |
853 |
854 |
855 |
856 |
857 |
858 |
859 |
860 |
861 |
862 |
863 |
864 |
865 |
866 |
867 |
868 |
869 |
870 |
871 |
872 |
873 |
874 |
875 |
876 |
877 |
878 |
879 |
880 |
881 |
882 |
883 |
884 |
885 |
886 |
887 |
888 |
889 |
890 |
891 |
892 |
893 |
894 |
895 |
896 |
897 |
898 |
899 |
900 |
901 |
902 |
903 |
904 |
905 |
906 |
907 |
908 |
909 |
910 |
911 |
912 |
913 |
914 |
915 |
916 |
917 |
918 |
919 |
920 |
921 |
922 |
923 |
924 |
925 |
926 |
927 |
928 |
929 |
930 |
931 |
932 |
933 |
934 |
935 |
936 |
937 |
938 |
939 |
940 |
941 |
942 |
943 |
944 |
945 |
946 |
947 |
948 |
949 |
950 |
951 |
952 |
953 |
954 |
955 |
956 |
957 |
958 |
959 |
960 |
961 |
962 |
963 |
964 |
965 |
966 |
967 |
968 |
969 |
970 |
971 |
972 |
973 |
974 |
975 |
976 |
977 |
978 |
979 |
980 |
981 |
982 |
983 |
984 |
985 |
986 |
987 |
988 |
989 |
990 |
991 |
992 |
993 |
994 |
995 |
996 |
997 |
998 |
999 |
1000 |

AAF81787 AAF81787 standard; cDNA; 2867 BP.

AC AAF81787;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:341.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; chromosome 1; ss.

OS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Kogen SM, Barash SC, Birse CE, Rosen CA;

PI WPI; 2001-235357/24.

DR P-PSDB; AAG73854.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -

PS Claim 1; Page 2452-2453; 9803bp; English.

XX AAF81787 to AAF81795 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAF81796 to AAF82204
CC and AAF81789 represent sequences used in the exemplification of the
CC present invention.

N.B. Pages 666 to 682 and page 7053 of the sequence listing were

used to diagnose, prevent and treat neoplastic, immunological and reproductive disorders. This sequence encodes a human membrane spanning protein MSP-5 fragment.

Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match 100.0%; Score 227; DB 20; Length 3111;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCCGGCTCTGGAGAGTAAAGCAGCCAGGAGATGATACCCAGAGAGTGA 60
DB 2056 GTGCGGCTCTGGAGAGTAAAGCAGCCAGGAGATGATACCCAGAGTGA 2115
QY 61 GCACGTACATACCTGGCTGGATGAGATGAATACATGATCCAGACACT 120
DB 2116 GCACTGTACATACCTGGCTGGATGAGATGAATACATGATCCAGACACT 2175
QY 121 GAAATTAATAGATGATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCTCGGGA 180
DB 2176 GAATTAATAGATGATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCTCGGGA 2235
QY 181 GGCTCATTTGGCTTGTGATGTCCTCAATGCTCCATACCTGATCT 227
DB 2236 GGCTCATTTGGCTTGTGATGTCCTCAATGCTCCATACCTGATCT 2282

RESULT 11

AAI29502
AAI29502 standard; cDNA; 3311 BP.

AAI29502;
12-OCT-2001 (first entry)

C902P determined cDNA sequence.

Human; immunotherapy; diagnosis; colon cancer; colon tumour;
Immunogenic; gene therapy; vaccine; colonic cancer; ss.

Homo sapiens.

WO200149716-A2.

12-JUL-2001.

29-DEC-2000; 2000MO-US35596.

30-DEC-1999; 9905-0476296.

10-JAN-2000; 2000US-0480321.

15-FEB-2000; 2000US-0504629.

06-MAR-2000; 2000US-0519444.

19-MAY-2000; 2000US-0575251.

29-JUN-2000; 2000US-0609448.

28-AUG-2000; 2000US-0649811.

(CORI-) CORIXA CORP.

XU J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;
King GE, Wang T, Jiang Y;

WPI; 2001-441847/47.

Colon tumor associated proteins and nucleic acids useful for the
prevention, diagnosis and treatment of colonic cancer

Claim 2; Page 425-426; 472pp; English.

The present invention describes colon tumour associated proteins (I) and
the polynucleotides (II) that encode them. (I) have cytostatic activity.
(I) and (II) can be used in gene therapy and vaccine production. (I) and
(II) may be used in the prevention, diagnosis and treatment of diseases

associated with inappropriate colon tumour associated protein (TCAP)
expression, such as colonic cancer. For example, (I) and (II) may be
used to treat disorders associated with decreased expression by
correcting mutations or deletions in a patient's genome to affect the
activity of TCAPs by expressing inactive proteins or to supplement the
patients own production of them. Additionally, (II) may be used to
produce the TCAP proteins, by inserting the nucleic acids into a host
cell culturing the cell to express the protein. (II) and its
complementary sequences may also be used as DNA probes in diagnostic
PCR polymerase chain reaction (PCR) and hybridisation assays to detect and
quantitate the presence of similar nucleic acids in samples, and
therefore which patients may be in need of restorative therapy. (I) may
also be used as antigens in the production of antibodies against TCAPs
and in assays to identify modulators of TCAP expression and activity.
Anti-(I) antibodies and antagonists may also be used to down regulate
TCAP expression and activity. The anti-(I) antibodies may also be used
as diagnostic agents for detecting the presence of TCAPs in samples
(e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
and AAI24494 to AAI24523 represent nucleotide and amino acid sequences
given in the exemplification of the present invention.

Sequence 3111 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;

Query Match 100.0%; Score 227; DB 22; Length 3311;
Best Local Similarity 100.0%; Pred. No. 1.1e-67;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCCGGCTCTGGAGAGTAAAGCAGCCAGGAGATGATACCCAGAGTGA 60
DB 2374 GTGCGGCTCTGGAGAGTAAAGCAGCCAGGAGATGATACCCAGAGTGA 2433
QY 61 GCACGTACATACCTGGCTGGATGAGATGAATACATGATCCAGACACT 120
DB 2434 GCACTGTACATACCTGGCTGGATGAGATGAATACATGATCCAGACACT 2493
QY 121 GAAATTAATAGATGATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCTCGGGA 180
DB 2494 GAATTAATAGATGATGTTCAACACAGCAAGTGTGTTTCAGCAGACATCTCGGGA 2553
QY 181 GGCTCATTTGGCTTGTGATGTCCTCAATGCTCCATACCTGATCT 227
DB 2594 GGCTCATTTGGCTTGTGATGTCCTCAATGCTCCATACCTGATCT 2600

RESULT 12

AAH46101
AAH46101 standard; DNA; 2739 BP.

AAH46101;
11-SEP-2001 (first entry)

Mouse Gob-5 coding sequence, SEQ ID NO:3.

Mouse Gob-5; murine; goblet cell; human C1CA1 orthologue; drug screening;
expression inhibition; antisense therapy; gene therapy; bronchial asthma;
chronic obstructive pulmonary disease; antiasmatic; ds.

Mus sp.

US

US

US

US

US

US

US

US

US

Location/Qualifiers

1..2739

/**tag- a

/product- "Mouse Gob-5"

/note- "No stop codon given in the specification"

MO300138530-A1.

31-MAY-2001.

22-NOV-2000; 2000MO-JP08232.


```

XX PR 24-NOV-1999; 99JP-0333479.
XX PR 27-APR-2000; 2000JP-0127589.
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX PI Nakanishi A, Morita S;
XX WP1; 2001-355935/37.
XX P-PSDB; AAB73715.
XX PT New antisense nucleotide, useful for treatment and prevention of
XX PT bronchial asthma and chronic obstructive pulmonary disease
XX PS Claim 3; Page 80-82; 104pp; Japanese.
CC CC The invention relates to an antisense nucleotide targeted to the mouse
CC CC Gob-5 gene (coding sequence shown in AAH46120) or its human counterpart,
CC CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC CC relates to an antibody specific for the Gob-5 protein, medical and
CC CC diagnostic compositions containing the antisense nucleotide or the
CC CC antibody, and methods and kits for screening for compounds which inhibit
CC CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC CC The antisense oligonucleotides and antibody are therefore useful for the
CC CC treatment and prevention of bronchial asthma and chronic obstructive
CC CC pulmonary disease. The present sequence represents the mouse Gob-5
CC CC gene coding sequence.
XX XX
XX SO Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other:
Query Match 46.2%; Score 104.8; DB 22; Length 2739;
Best Local Similarity 73.0%; Pred. No. 9e-26;
Matches 165; Conservative 0; Mismatches 52; Indels 9; Gaps 2;
OY 5 GGGCTCTGGAGAGATTAAACGACCCACAGACAGACTGATACCCACAGAGAGTGAGCAGC 64
Db 2030 gggctctggagagattaaacgacccacagacagactgataccacagagagtggagcac 2089
OY 65 TGTACATCACTGGCGGATTTGAAATGATGAATCAATGAAATCCAGACCAAGCCTGAAA 124
Db 2090 tgtacatcactggcggaattgaaatgatgaatcaatgaaatccagaccagcttgaaa 2149
OY 125 TTAATAGGATGATGTTCAACACAAAGCAAGTGTGTTTCAGCGAAGCATCTCGGAGCGT 184
Db 2150 cta-----gtatgttcaagacaagacgctgtgttcacagacatctcaaggaggat 2203
OY 185 CATTTGTGGCTTCGATGTC---CAAAATGCTCCCATACCTGATCAT 227
Db 2204 cgttgtggccaccaatgtcccgacagcagctcccatctcgaact 2249
RESULT 13
AAH46120
XX ID AAH46120 standard; cDNA; 2843 BP.
XX AC AAH46120;
XX DT 11-SEP-2001 (first entry)
XX DE Mouse Gob-5 cDNA, SEQ ID NO:22.
XX OS Mus sp.
XX PH Key Location/Qualifiers
XX FT CDS 15..2756
XX FT /*tag= a
XX FT /product= "Mouse Gob-5"
XX FT

```

```

01 W06000138530-A1.
02
03
04 31 MAY-2001.
05
06 XX
07 FE 22 NOV-2000; 2000OWO-JP08232.
08 XX
09 PR 24 NOV-1999; 99JP-0333479.
10 PR 27 APR-2000; 2000UP-0127589.
11 PA (TAKE ) TAKEDA CHEM IND LTD.
12
13 XI Nakanishi A, Morita S;
14 PI WPL: 2001-355935/37.
15 DR P-PSDB; AAB73715.
16 XX New antisense nucleotide, useful for treatment and prevention of
17 XX bronchial asthma and chronic obstructive pulmonary disease -
18 XX Example 1; Page 89-91; 104pp; Japanese.
19
20 CC The invention relates to an antisense nucleotide targeted to the mouse
21 CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
22 CC the CLICAL gene (coding sequence shown in AAH46102). The invention also
23 CC relates to an antibody specific for the Gob-5 protein, medical and
24 CC diagnostic compositions containing the antisense nucleotide or the
25 CC antibody, and methods and kits for screening for compounds which inhibit
26 CC the protein. Gob-5 and CLICAL are proteins expressed by goblet cells.
27 CC The antisense oligonucleotides and antibody are therefore useful for the
28 CC treatment and prevention of bronchial asthma and chronic obstructive
29 CC pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
30
31 SQ Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;
32
33 Query Match 46.2%; Score 104.8; DB 22; Length 2843;
34 Best Local Similarity 73.0%; Pred. No. 9,1e-26;
35 Matches 165; Conservative 0; Mismatches 52; Indels 9; Gaps 2
36
37 QY 15 GGCGCTGTGGAGGAGATTAAACGAAGCCAGCAGAGATGATACCAGCAAGTGGAGCAC 64
38 | ||||| ||||| ||| ||||| ||||| ||| |||||
39 Db 2044 gggctgtgaggaaggatcacttcagacagagaagcaagcatcccgaaagacaagcca 2103
40
41 QY 15 TGTCACATACTTGCGCGCATTTGAGAATGATGAATACCAATGCAGATCCACCAAGCCGGAAA 124
42 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
43 Db 2104 tgcatactaaggcuggatttgaggagtggtaagatlaagaatgccaccacagttcgaata 2163
44
45 QY 135 TTAAATPACGATGATGTTCAACAACAAGCAAGTAGTGTTTTTCAGACGAACATCCTCGGAGGCT 184
46 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
47 Db 2164 cea-----gttatgttcaaagaacagcagctgtgtcctcagcagacatcttcagggygat 2217
48
49 QY 185 CATTTGTGGCTTCTGATGCC---CAAATGCTCCCCTACTACTATCT 227
50 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 Db 2218 cglttgtgccaccaatgtcccgcgcagcagctcccatctctgacct 2263
52
53 RESULT #14
54 AAF81925
55 AC XA F81925 standard, cDNA; 2931 BP.
56 XX
57 XX AAF81925;
58
59 DE 13 JUN-2001 (first entry)
60
61 PE Mouse ICACC-1 nucleotide sequence.
62
63 KW ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
64 KW interleukin 9 induced calcium activated chloride channel; IL-9;
65 KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
66 KW anti-inflammatory; immunomodulatory; cystic fibrosis;
67 KW inflammatory bowel disease; autoimmune disease; ss.
68
69 XX
70 XX Mus sp.
```


Thu Apr 4 09:27:28 2002

Db 1669 ggtgatttgggtatcaccaagtcaccaagccttccttgctgacc 1714

Search completed: April 4, 2002, 03:18:15
Job time: 60247 sec

us-09-049-696-13.rms

Page 10

(uspto) Page Blank
This is

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:03:29 ; Search time 22700.8 Seconds
(without alignments)
164.966 Million cell updates/sec

Title: US-09-049-696-13

Perfect score: 227
Sequence: 1 GTGGCGGCTCTGGGAGGAGT.....AATGCTCCATACCTGATCT 227

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Genbankl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rod: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	227	100.0	878	6	195746	195746 Sequence 8
2	227	100.0	2022	9	AK024970	AK024970 Homo sapi
3	227	100.0	2826	9	AF127036	AF127036 Homo sapi
4	227	100.0	3311	6	AX193489	AX193489 Sequence
5	227	100.0	3311	9	AF039400	AF039400 Homo sapi
6	142.2	62.6	3079	4	AF095584	AF095584 Sus scrofa
7	137	60.4	35278	4	AF039401	AF039401 Homo sapi
8	137	60.4	113764	9	HS3651E10	HS3651E10 Homo sapi
9	137	60.4	164891	2	AL358950	AL358950 Homo sapi
10	104.8	46.2	2937	10	AB017156	AB017156 Homo sapi
11	94.8	41.8	1895	9	AK000138	AK000138 Homo sapi
12	94.8	41.8	3304	9	AF127035	AF127035 Homo sapi
13	94.8	41.8	3321	9	AK000072	AK000072 Homo sapi
14	94.8	41.8	3265	6	AK092338	AK092338 Sequence
15	72.2	31.8	2984	4	BT036445	BT036445 Bos taurus
16	64.2	28.3	2820	4	AF001263	AF001263 Bos taurus
17	64.2	28.3	3288	4	AF001262	AF001262 Bos taurus
18	64.2	28.3	3317	4	AF001261	AF001261 Bos taurus
19	61.2	27.0	2765	10	AF115852	AF115852 Mus muscu
20	61.2	27.0	3137	10	BC008147	BC008147 Mus muscu
21	59.6	26.3	3058	10	AF108501	AF108501 Mus muscu
22	54.8	24.1	3022	10	AF047838	AF047838 Mus muscu
23	54.8	24.1	3471	10	AF052746	AF052746 Mus muscu
24	53.8	23.7	3415	9	AF043976	AF043976 Homo sapi
25	53.2	23.4	140718	2	AL356270	AL356270 Homo sapi
26	38.2	16.8	213673	2	AL593846	AL593846 Mus muscu
27	36	15.9	119657	2	AC073667	AC073667 Mus muscu
28	36	15.9	208979	2	AC023610	AC023610 Mus muscu
29	35.4	15.6	91351	9	AL158850	AL158850 Human DNA
30	35.4	15.6	173286	2	AL592434	AL592434 Homo sapi
31	34.6	15.2	129503	2	AC044863	AC044863 Mus muscu
32	34.4	15.2	157141	9	AC011470	AC011470 Homo sapi
33	34.4	15.2	167509	2	AC012107	AC012107 Homo sapi
34	34.2	15.1	23005	2	AC024888	AC024888 Homo sapi
35	34.2	15.1	242159	2	AC091567	AC091567 Homo sapi
36	34	15.0	66736	9	HS262817	HS262817 Human DNA
37	34	15.0	99342	9	HS201D17	HS201D17 Human DNA
38	34	15.0	136693	9	HS16483	HS16483 Human DNA
39	33.8	14.9	130988	9	HS364122	HS364122 Human DNA
40	33.8	14.9	148468	2	AL355882	AL355882 Homo sapi
41	33.8	14.9	165599	2	AL353740	AL353740 Homo sapi
42	33.8	14.9	193492	2	AL590072	AL590072 Homo sapi
43	33.6	14.8	116210	9	HS406C18	HS406C18 Human DNA
44	33.6	14.8	149341	2	AC068451	AC068451 Homo sapi
45	33.6	14.8	175591	2	AC068071	AC068071 Homo sapi

ALIGNMENTS

RESULT: 1
LOCUS 195746 878 bp DNA
DEFINITION Sequence 8 from patent US 5733748.
ACCESSION 195746
VERSION 195746.1 GI:3940216
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 878)
AUTHORS Yu, G. and Rosen, C.
TITLE Colon specific genes and proteins
JOURNAL Patent: US 5733748-A 8 31-MAR-1998;
FEATURES
source 1. 878
BASE COUNT 257 a 179 c 188 g 241 t 13 others
ORIGIN

Query Match 100.0%; Score 227; DB 6; Length 878;
 Best Local Similarity 100.0%; Pred. No. 1,4e-61;
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGGGCTCTGGAGAGTTAACGACAGCAGAGAGATATACCCAGCAGAGTGA 60
 DB 56 GTGGGGGCTCTGGAGAGTTAACGACAGCAGAGAGATATACCCAGCAGAGTGA 115
 QY 61 GCAGTGCATACCTGGCTGGATGAGATGAAATACAAATGCAATCCACCAAGACT 120
 DB 116 GCAGTGCATACCTGGCTGGATGAGATGAAATACAAATGCAATCCACCAAGACT 175
 QY 121 GAAATTAATAGATGATGTTCAACACAGCAAGTGTGTTAGACAGAAATCCTCGGGA 180
 DB 176 GAAATTAATAGATGATGTTCAACACAGCAAGTGTGTTAGACAGAAATCCTCGGGA 235
 QY 181 GGCTCATTTGTGGCTTGTGATGTCACCAATGCTCCATACCTGATCT 227
 DB 236 GGCTCATTTGTGGCTTGTGATGTCACCAATGCTCCATACCTGATCT 282

RESULT 2

AK024970 2022 bp mRNA PRI 29-SEP-2000
 LOCUS AK024970 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA.
 DEFINITION Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA.

ACCESSION AK024970.1 GI:10437397
 VERSION AK024970.1
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens colon cDNA to mRNA, clone_11b:COL clone:COL02275.
 ORGANISM Homo sapiens

REFERENCE 1 (sites)
 Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 2 (bases 1 to 2022)
 Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T. and Nakamura, Y.
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology: cDNA library construction, 5'-6' and one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source location/Qualifiers
 1..2022
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COL02275"
 /clone_11b="COL"
 /tissue_type="colon"
 /note="cloning vector pME18SFL3"
 misc-feature
 1..2022
 /note="highly similar to AF127036 Homo sapiens calcium-activated channel protein 1 (CaCC1) mRNA"
 BASE COUNT 612 a 472 c 453 g 485 t

Query Match 100.0%; Score 227; DB 9; Length 2022;
 Best Local Similarity 100.0%; Pred. No. 1.6e-61;
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGGGCTCTGGAGAGTTAACGACAGCAGAGAGATATACCCAGCAGAGTGA 60
 DB 1236 GTGGGGGCTCTGGAGAGTTAACGACAGCAGAGAGATATACCCAGCAGAGTGA 1295
 QY 51 GCAGTGCATACCTGGCTGGATGAGATGAAATACAAATGCAATCCACCAAGACT 120
 DB 1236 GCAGTGCATACCTGGCTGGATGAGATGAAATACAAATGCAATCCACCAAGACT 1355
 QY 121 GAAATTAATAGATGATGTTCAACACAGCAAGTGTGTTAGACAGAAATCCTCGGGA 180
 DB 1356 GAAATTAATAGATGATGTTCAACACAGCAAGTGTGTTAGACAGAAATCCTCGGGA 1415
 QY 181 GGCTCATTTGTGGCTTGTGATGTCACCAATGCTCCATACCTGATCT 227
 DB 1416 GGCTCATTTGTGGCTTGTGATGTCACCAATGCTCCATACCTGATCT 1462

RESULT 3

AF127036 2826 bp mRNA PRI 10-AUG-1999
 LOCUS AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA, complete cds.
 DEFINITION AF127036
 ACCESSION AF127036
 VERSION AF127036.1 GI:4585468
 KEYWORDS human.
 SOURCE Homo sapiens

REFERENCE 1 (bases 1 to 2826)
 Agnel, M., Verma, T. and Culouscou, J.-M.
 IDENTIFICATION of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea
 FEBS Lett. 455 (3), 295-301 (1999)

REFERENCE 2 (bases 1 to 2826)
 Agnel, M. and Culouscou, J.-M.
 Direct Submission
 Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrieres, Reuil-Malmaison 92500, France
 JOURNAL Location/Qualifiers
 1..2826
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="small intestine; colon"
 /product="calcium-activated chloride channel protein 1"
 /protein_id="A4D25487.1"
 /db_xref="GI:4585469"

FEATURES

gene
 CDS
 1..2826
 /gene="CaCC1"
 /note="bovine epithelial chloride channel homolog"
 /codon_start=1
 /product="calcium-activated chloride channel protein 1"
 /protein_id="A4D25487.1"
 /db_xref="GI:4585469"
 /translation="MGPKSSVFLILHLEAGALSNSLIQINNNYEGIVADPNP
 EDENLQIKDMVQASLYLPEFGKRGKRGKRVNVALILPEFKTADYVPRLEYKNA
 DVAIAESTPPRENDERPYROMNCGEKGRIHLIPDYLAKKTLAYGPGRAFYEMH
 LRMGVEDYNDKRYLSNGRIQAVRCGAGITGRVYKKGCGSCYTRKCTFNKVTY
 LKGCDFYLSRQTEKASIMPAQVDSIVERTQONHKEAPKQNCNIRSTWEY
 RDSDFKFTPTMTQPPNPTSLQIGRIYCLVDRSGSMATNRNRLRNQAGLFL
 LQYVLESGMWGAVTFDSAAHYQSELQINSQSDRTYAKRLPAAASGTSQCELSRA
 FTVIRKRYPTDGESEVILLTDEDMTISGCFNEVROSGAIIHTVALGSAOCELELSK
 MTCGLQYASQVQVONGLIDAPGALSNGVAVSQRSGGCFVYDKNTNATYIQLPQAKVETWY
 DSTVAKDLFLITWTTPQILMDPSQKQKGFVYDKNTNATYIQLPQAKVETWY
 SIQASSQFLTLIVYSRASNATLPPTVYSKTKNKTQSKPSPPLVYANRQASPIIIRA
 SVTALIESVNGKTVYIELLDNGAGADATKDGVSRYFTYDVTNGRRSVKVRALGVAN
 AARRVYIQSGALYIPGMIEDEIOWMPPEIPKIDVDHQRQVCFSTSGSGSFVAS

CDS

/gene="hclcal1"
/number=2
join(2222..2383,6667..6807,6979..7126,9696..9801,
15453..15631,18591..18809,19783..20010,22253..22427,
24522..24628,26644..26859,27447..27708,28766..28936,
31840..32079,32919..33310)
/gene="hclcal1"
/product=start-1
/codon_start=1
/protein_id="AAC95429.1"
/db_xref="GI:4009460"
/translation="MGPFKSSVFLLILHLEGLALSNSLIQIINNNGYEGIVVAILDPNVP
EDELLIOIKDMVTQASLYLFEAIGKRFRFKNVAIIIPETMKTADYVRPLETEYRNA
DVALVEASTPGNDPEPYEOMGNGEKGERIHLPDPITAGKKIAEYGPGAKAFHEMAH
LRNGPDEVNNDEKFVLSNGRIQAVRCSAGITGNVYKKGCGSCYTRCTFPNTGTL
YKCECFEVLDSROTTERKASIMPAQHVDISIVEFCFBONHNKAEPKONCKIRNSTWEIT
RDSDSEFKTTPMTTPPPNPFFSLIQICQRTVCVLVDKSGSMATGNRNRLNRINQAQLEFL
LQYVELGSWGMVTFDSAHVQSELIDINGSDSDTLAKRLPAASGSTISCSGLRSBA
FTVYARKYPYDGSEIVYLTDGEDNTISGCNEVSDGAIITHVALGSAOEELELSK
MTGGLOIYASDOVQNNGLIDAFGALSGNGAVSORSIOLESKGLTIQNSOMNGTVIIV
DTPTGKDITFLITWTQPOIILMDPSGOKGVVVKNKRMAYIOLPIGLAKVCTMKY
SIQASSOTLTJVTYSRASNAITPIYTSTKTNKDESKFPSPIVVYANNROGASPIILA
SVAILRESQKVYTELELDNGACADA TKDGYSRIFTTYDNGRSIVYRAALGVNV
AARRRVIPQOSGALYIPGWIENEDIQNNPRELNKDDVOHKOVCSFSVSGSFVAS
LVDPNAPIPDLFPEGQITDLNABKEHGSLINLTWPAPDGDGHGAHYIIRISISLD
DRDFNESILOVNTFALIIPKANSEEVLFPENTFENGSDLFAIQAIVKVIDKTSKI
SNIAARVSLFIPTQPPEPTSPDETSACPNIHINSTIPGIHLKIMKMWIGELQSLIA
"

exon

6667..6807
/gene="hclcal1"
/number=3
6979..7126
/gene="hclcal1"
/number=4
9696..9801
/gene="hclcal1"
/number=5
15454..15631
/gene="hclcal1"
/number=6
18591..18809
/gene="hclcal1"
/number=7
19783..20010
/gene="hclcal1"
/number=8
22253..22427
/gene="hclcal1"
/number=9
24522..24628
/gene="hclcal1"
/number=10
26644..26859
/gene="hclcal1"
/number=11
27447..27708
/gene="hclcal1"
/number=12
28766..28936
/gene="hclcal1"
/number=13
31840..32079
/gene="hclcal1"
/number=14
32919..33322
/gene="hclcal1"
/number=15

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

BASE COUNT

11245 a 6870 c 7230 g 9933 t

ORIGIN

Query Match	60.4%;	Score 137;	DB 9;	Length 35278;
Best Local Similarity	100.0%;	Pred. No. 1e-32;		

```

Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 91 GATCAATACATGATGATCCACCAAGACCTGTAATTAATAGATGATGTTCAACAAAG 150
|||||
CY 31819 GATGAATATACATGAATGAATCCACCAAGACCTGTAATTAATAGATGATGTTCAACAAAG 31898
|||||
CY 151 CAAGTGTTCACAGAGAACATCTCGGAGAGCTATTGGCTTCGATGATGCCAAT 210
|||||
DB 31819 CAAGTGTTCACAGAGAACATCTCGGAGAGCTATTGGCTTCGATGATGCCAAT 31958
|||||
CY 221 GCTCCATACCTGATCT 227
|||||
DB 31919 GCTCCATACCTGATCT 31975
|||||

RESULT 8
LOCUS HSJ651E10
DEFINITION Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1, complete sequence.
ACCESSION AL122002
VERSION AL122002.16 GI:8247274
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 113764)
AUTHORS Brown,A.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
REQUESTS: clonequests@sanger.ac.uk
ON Jun 4, 2000 this sequence version replaced gi:8247030.
During difference assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unseq' feature key.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr1
RP4-651E10 is from the library RPhi-4 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP4-651E10.
FEATURES
Location/Qualifiers
1..113764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p22.3-31.1"
/clone="RP4-651E10"
/clone_id="RPhi-4"
BASE COUNT 34890 a 21989 c 22351 g 34534 t
ORIGIN
Query/Match 60.4%; Score 137; DB 9; Length 113764;

```


DEFINITION Mus musculus gob-5 mRNA, complete cds.
 ACCESSION AB017156
 VERSION AB017156.1 GI:3721911
 KEYWORDS gob-5.
 SOURCE Mus musculus adult intestine goblet cell cDNA to mRNA.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (sites)
 AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.
 TITLE Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice
 JOURNAL Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)
 MEDLINE 99160866
 REFERENCE 2 (bases 1 to 2937)
 AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.
 TITLE Direct Submission
 JOURNAL Submitted (26-AUG-1998) to the DDBJ/EMBL/GenBank databases. Tohru Komiyama, ERIYO, JST, Genome Asymmetry Group, Doi Biosymmetry Project, Tokodai 5-9-9, Tsukuba, Ibaraki 300-2635, Japan (E-mail:tkomedia@erauto.tirc-net.co.jp, Tel:81-298-48-1515, Fax:81-298-47-8901)

FEATURES
 source
 1..2937
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_type="goblet cell"
 /dev_stage="adult"
 /tissue_type="intestine"
 15..2756
 /gene="gob-5"
 15..2756
 /gene="gob-5"
 /codon_start=1
 /protein_id="BA33743.1"
 /db_xref="GI:3721912"
 /translation="MESLSPVFLILHLLEGVLSLSLIQNNNGYEGIVIAIDHVP
 EDALQIKDMVTOASPYLFEGATGKRYFKVVALIPESWAKPEYTPKLETKNA
 DLVSTPTPLKNDPEYTHGACGKGIHLTPPLGKRLTOYGPORDTPVHEAH
 FRMGVENEYNDKEFYLSKGRQAVRCSALITGKNOVRGCGSCITNGKVIDRGTG
 LYKNCVFPDPHONERKASIMNONINSVECTPKNHNOEAPNNOONCNLRSTREY
 IOESDEKOTTPMTAPAPPTSLQIIGRTICVLDSGSLNDRLRNMQASFLF
 LLQIVEGSWGMVTFDSAAVQSEIKLNSGADBDLILKLPVSAAGTSICSGRT
 AFTVTKRYPDGESEIVLTDDEMTISSCPDVOSGAIHTVALGPAALAEGLDS
 KMTGGLQVYSPOVONNGLVDAFALSSGNALIAHISIOLESRVNIAONNOMNSGI
 VDSVSGDPLFLITWTHTPPTFEIMDPSGVONGILDTTTRKAVILOYGTAKVGMK
 LRDHNTSLQVNTTGLIPKEASSESEIFEFELGNTFNGDIFIAIOAVDKSNLSEI
 SNIAVSVEIFPAOEPIPEDSPPCPDISINSTIPGIIHVKIMMKWLGDMQVTLGH"
 BASE COUNT 860 a 718 c 693 g 666 t
 ORIGIN

Query Match 46.2%; Score 104.8; DB 10; Length 2937;
 Best Local Similarity 73.0%; Pred. No. 1.4e-22;
 Matches 165; Conservative 0; Mismatches 52; Indels 9; Gaps 2;

DB 5 GGGCTCTGGAGAGGTTAAACGACGACGAGAGTATACCCCGACGAGATGAGCAC 64
 |||||||
 DB 2044 GGGCTCTGGAGAGGTTAAACGACGACGAGAGTATACCCCGACGAGATGAGCAC 2103
 |||||||
 DB 65 TGTACATACCTGCTGGATTTGAGATGATGAATACATGATGATCCACCAAGACCTGA 124
 |||||||
 DB 2104 TGTACATACCTGCTGGATTTGAGATGATGAATACATGATGATCCACCAAGACCTGA 2163
 |||||||
 DB 125 TTAATAGAGTATGTTTCAACACCAAGAGTGTGTTTCACGACAGACATCCTCGGAGAGCT 184
 |||||||
 DB 2164 CTA-----GTTATGTTCAAGACACGAGCTGTGCTTCAAGACAGACATGTTCAAGGGGAGT 2217
 |||||||

RESULT 11
 AK000138
 LOCUS Homo sapiens cDNA FLJ20131 fis, clone COL06357.
 DEFINITION AK000138
 ACCESSION AK000138.1 GI:7020030
 VERSION
 KEYWORDS oligocapping; fis (full insert sequence).
 SOURCE Homo sapiens colon cDNA to mRNA, clone 11b:COL clone:COL06357.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Odayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished (2000)
 REFERENCE 2 (bases 1 to 1895)
 AUTHORS Sugano, S., Suzuki, Y., Ota, T., Odayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.
 TITLE Direct Submission
 JOURNAL Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ems.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'-3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES
 source
 1..1895
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COL06357"
 /clone_11b="COL"
 /tissue_type="colon"
 /note="cloning vector pME18SFL3"
 66..1475
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAA90969.1"
 /db_xref="GI:7020031"
 /translation="MSKITGSGSHFYVSDAONNGLIDAFGALTSGNTDLSOKSLAES
 KGLITNSMAMNDVTIIDSTGKDPFELITWNPSISLIMPSCGIMENFTYDASK
 MAVLSIRCTAKYCTAAYNIQAKANPELTITITYSRAANSVPPIYTNAMKNDVNSFP
 SPMTVAETLQGVNVAFTAFIESQNGHTEVLELLDNGKADSFNDGYSRIFT
 ATTENGRYSLKVARHAGANTARLKLPLNRAAYIPGWVNEIENEMPREPEIDETO
 TLEDEFSTRDAGFAVSOVPSLPDPQYPSQIDLVADIVHEDKILTWTFAGDNED
 VGKVOYIIRISASIIDLRSDDDALQVWTTPLDSPEANSKESFAKPNISSENAETH
 IFPAISIDKSNITSKVSNIQAQVTLPIQANDPDIDPPTPTPDKSHNSGVNISTLV
 VLSVSGSVIVNPIILSTT"
 BASE COUNT 658 a 352 c 357 g 528 t
 ORIGIN

Query Match 41.8%; Score 94.8; DB 9; Length 1895;
 Best Local Similarity 66.8%; Pred. No. 2.1e-19;
 Matches 151; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

DB 1 GTCGGGCTCTGGAGAGGTTAAACGACGACGAGAGTATACCCCGACGAGAGTGA 60
 |||||||
 DB 753 GTTCGGGCTCTGGAGAGGTTAAACGACGAGAGTATACCCCGACGAGAGTGA 812
 |||||||
 DB 31 GCACTGACATACCTGCTGGATTTGAGATGATGAATACATGGAATCCACCAAGACT 120
 |||||||

Db 813 GCCGCTACATACAGCGCTGGTAGTGNACGGGAAATGAGCAAAACCCGCAAGACT 872

Qy 121 GAAATTAATAGGATGCTTCAACACAGCAAGAGTGTCTTACAGCAACATCTCCGGGA 180

Db 873 GAAATTTGAT---GAGCATACTACAGCACCTTTGAGAGATTTTACAGCCGAAAGCATCCGGA 929

Qy 181 GGCCTATTTTGGCTTCTGATGTGCCAAATGCTCCCATCTGATC 226

Db 930 GGTGCATTTGTGTATCAACAAGTCCCAAGCCTTCCCTTGCCGACC 975

RESULT 12

AF127035 3204 bp mRNA PRI 11-AUG-1999

LOCUS Homo sapiens calcium-activated chloride channel protein 2 (CaCC2)

DEFINITION mRNA, complete cds.

ACCESSION AF127035

VERSION AF127035.1 GI:5726288

KEYWORDS

ORGANISM human.

REFERENCE 1 (bases 1 to 3204)

Agnel, M., Vernat, T., and Culouscou, J. M.

Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea

FEBS Lett. 455 (3), 295-301 (1999)

JOURNAL MEDLINE 99364503

PUBMED 10437792

REFERENCE 2 (bases 1 to 3204)

Agnel, M., and Culouscou, J. M.

Direct Submission

Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrieres, Ruell-Malmaison 92500, France

JOURNAL FEATURES

source Location/Qualifiers

1..3204

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="colon"

1..3204

/gene="CaCC2"

29..2782

/note="CaCC2"

/note="bovine epithelial chloride channel homolog"

/codon_start=1

/product="calcium activated chloride channel protein 2"

/protein_id="MAD48398.1"

/db_xref="GI:5726289"

/translation="MGLFRGFVFLVLCILHOSNTSFIKLNNGEDIVIVIDPSVE DEKTIIOIDMTTASTLYLEATEKEEFKKNVSLIPENMKENPQYKRPKHENKHAH VIVAPPTLGRDREPPYKOPTEGEKEGEYIHPDPLLGKKNQYPPGKLFYEMAHL RMGVEDEINEDOPFYRAKSKIKATKCSAGISGRNRYVYCCGSCSRACRIDSTKTL YGKDCQFFPDQTERKASIMFQSIDSVVEFCERKHNOBAPLONIKRNPSTWVI SNSEDKNTIPMTVPPPPVSLKTRORIVCLVDRSSMGKRLNNOAKHFL LOTVENGSMGAVHFDSTATIVNKLQIKSDERNIMAGLPPYPLGSGISGKYA FOYIGELHSDGSEVLLIDGENDTASSCIDVKSOGAIVHFIAGRADAVIEMS KITGSHFYVSDRANONGIIDAFALGSGNTLSOKLSKGLTNLANMAMNDVI IDSTVGKDTFFLTMSNPSPISIMPSGCTIMNTVDAATSKAVLIPGTAKVGTWA YNLQAKNPETLITVTSRANSSVPPIYNAKMDVNSPSPMIVAVETILGQYAV LGANVTAFTESONGHTVEVLLDNGAGDSFKNDGYSTRFTIYTENGRYSLKRAHG GANTARLKLPLNRAVYIPGVVNGEIEANPREPIDEIDTDTLDEFSKSGAAY VSOYSLPDPDPSQITDLDAIVHEDKILITWTAGDNFDGKYQRYIIRISAIL DLRSPDADLOVMTTDSKPEANSKESFAKPNISSENAITHIFAIKSIDKSNLTSK VSNIAQVTLTLPQANPDIDPTPTPTPKSKNSGVNISTLVLSVIGSVYIVNFILS TTT

BASE COUNT 1098 a 594 c 633 g 879 t

ORIGIN

Query Match 41.8%; Score 94.8; DB 9; Length 3204;

Best Local Similarity 66.8%; Pred. No. 2.2e-19;

Matches 151; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

Qy 121 GTGGGGCTCTGGAGAGCTTAACGACGACGAGAGAGTATACCCAGAGAGTGA 60

Db 260 GTTCGGGCTCATGAGAGAGAGAAACACTGCCAGCTAAATTTACGGCTTCAGTAAGA 2119

Qy 61 GCACCTFACATACCTGCTGATTTGAGAAATGATGAATACATGAAATCCACCAAGCT 120

Db 2120 GCCGCGTACATACAGAGCTGGTAGTGAACGGGAAATTTAAACCAACCCGCAAGACT 2179

Qy 121 GAAATTAATAGGATGCTTCAACACAGCAAGAGTGTCTTACAGCAACATCTCCGGGA 180

Db 2180 GAAATTTGAT---GAGCATACTACAGCACCTTTGAGAGATTTTACAGCCGAAAGCATCCGGA 2236

Qy 181 GGCCTATTTTGGCTTCTGATGTGCCAAATGCTCCCATCTGATC 226

Db 2237 GGTGCATTTGTGTATCAACAAGTCCCAAGCCTTCCCTTGCCGACC 2282

RESULT 13

AK000072 3221 bp mRNA PRI 22-FEB-2000

LOCUS Homo sapiens cDNA FLJ20065 fis, clone COL01613, highly similar to ECDC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.

DEFINITION

ACCESSION AK000072

VERSION AK000072.1 GI:7019922

KEYWORDS oligo capping; fis (full insert sequence).

REFERENCE Homo sapiens colon cDNA to mRNA, clone_11b:COL clone:COL01613.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3221)

Sugano, S., Suzuki, Y., Ota, T., Odayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T., and Nakamura, Y.

Direct Submission

Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

JOURNAL COMMENT

NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'-6' and one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source Location/Qualifiers

1..3221

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="COL01613"

/clone_11b="COL"

/tissue_type="colon"

/note="cloning vector pME18SF13"

1..3221

/note="highly similar to ECDC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"

BASE COUNT 1105 a 600 c 634 g 882 t

ORIGIN

Query Match 41.8%; Score 94.8; DB 9; Length 3221;

Best Local Similarity 66.8%; Pred. No. 2.2e-19;

Matches 151; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

Qy 1 GTGGGGCTCTGGAGAGCTTAACGACGACGAGAGAGTATACCCAGAGAGTGA 60

Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:19 ; Search time 16681 Seconds

(without alignments)
162.336 Million cell updates/sec

Title: US-09-049-696-6

Perfect score: 252
Sequence: 1 CAAGAGATTGTGTGTAGT.....GAGGAGCTCATCTGCAGC 252

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estlov:*
9: em_hic:*
10: qb_est1:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_fod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	248.8	98.7	435	10	AM844161 RC4-CN004
2	169	67.1	460	10	AM750670 CM2-CN003
3	164.2	65.2	2933	12	AK008659 Mus muscu
4	162.6	64.5	2915	12	AK007466 Mus muscu
5	160.6	63.7	704	11	BG965023 602831668
6	157.4	62.5	890	11	BF578833 602094876
7	140.4	55.7	327	11	FI5082 SSG6A10 Por
8	110	43.7	831	11	BG286204 602383026
9	107.6	42.7	640	11	BF320704 uz55f02.Y
10	107.4	42.6	453	11	AM258167 uq3ic11.Y
11	105.8	42.0	1047	11	BG247065 602359723
12	99.8	39.6	949	11	BF099712 601751347

13	99.2	39.4	372	10	AA755896 vu04f06.f
14	97.2	38.6	402	10	BE656570 UI-M-BHO-
15	86.8	34.4	480	10	AA581198 nd38c07.r
16	81.2	32.2	661	11	BF581041 602100173
17	76	30.2	972	11	BG173173 602335553
18	71	28.2	208	10	BS587551 BS587551
19	70.8	28.1	747	11	BF580957 BF580957
20	68.4	27.1	433	11	BE916629 602100659
21	63.4	25.2	1016	11	BF102225 601752527
22	55.6	22.1	156	13	BF102225 601752527
23	53.8	21.3	823	11	AB49747 2M0151116
24	53.8	21.3	894	11	BG741082 602631626
25	48	19.0	305	10	BG677576 602624129
26	47.8	19.0	1033	11	BE240298 BE240298
27	41.6	16.5	858	11	BG243305 602355738
28	40.2	16.0	563	13	BG675746 602621972
29	39.4	15.6	602	10	A0437882 HS_5070.A
30	37	14.7	118	10	AL602490 DKFZP686P
31	36.6	14.5	307	10	AM841718 RC3-CN001
32	34	13.5	881	13	AM577823 CM0-CN000
33	33.6	13.3	410	11	A2184131 SP_1002.B
34	33.2	13.2	869	11	BF960289 QV2-NN004
35	33	13.1	338	11	BI086311 602849540
36	33	13.1	875	11	R01958 yeb5c07.r1
37	32.8	13.0	548	13	BF316958 601803911
38	32.6	12.9	455	11	AO790256 HS_2240.B
39	32.6	12.9	537	11	BI253970 602974391
40	32.6	12.9	835	11	BI045693 MR3-FN020
41	32.4	12.9	840	11	BG309020 HVSMEC000
42	32.4	12.8	301	13	BF531441 602091546
43	32.2	12.8	301	13	BF531441 602091546
44	32.2	12.7	739	11	BH039199 RPI-24-2
45	31.8	12.6	552	13	BH029381 RPI-24-2
					BF960643 QV2-NN004
					AZ356696 IM0097F21

ALIGNMENTS

RESULT 1
LOCUS AM844161 435 bp mRNA
DEFINITION RC4-CN0048-150400-013-f07 CN0048 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM844161
VERSION AM844161.1 GI:7938144
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Mitsuikuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL
MEDLINE
COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?rl=62-RC4-CN0048-150400-013-f07&3=2000-04-15&4=1)
Seq primer: puc 18 forward

High quality sequence start: 63
High quality sequence stop: 434.
Location/Qualifiers

FEATURES

source

1. 435
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CN0048"
/dev_stage="Adult"
/note="Organ: colon, normal; Vector: puc18; Site: 1; SmaI; Site: 2; SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 124 a 105 c 115 g 91 t
ORIGIN

Query Match 98.7%; Score 248.8; DB 10; Length 435;
Best Local Similarity 99.2%; Pred. No. 4,7e-66;
Matches 250; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAAGAATGTGTGTAGTCTTGTGACAAATCTGAGACATGGCGACTGTACCGCTC 60
|||||
Db 111 CAAGAATGTGTGTAGTCTTGTGACAAATCTGAGACATGGCGACTGTACCGCTC 170
OY 61 AATGCACTGAATCAAGCAGCCAGCTTTCTGCTGACAGACAGTTGAGTGGGCTCTG 120
|||||
Db 171 AATGCACTGAATCAAGCAGCCAGCTTTCTGCTGACAGACAGTTGAGTGGGCTCTG 230
OY 121 GTTGGAGTGTGACATTTGACAGTGTGCTGCCATGTACAAAGTGAACATCAATTAAC 180
|||||
Db 231 GTTGGAGTGTGACATTTGACAGTGTGCTGCCATGTACAAAGTGAACATCAATTAAC 290
OY 181 AGTGCAGTGAAGCAGGACACACTCCGCCAAAGATTACTGTCAGACGCTTCAGAGAGGAGC 240
|||||
Db 291 AGTGCAGTGAAGCAGGACACACTCCGCCAAAGATTACTGTCAGACGCTTCAGAGAGGAGC 350
OY 241 TCCATCTGCAGC 252
|||||
Db 351 TCCATCTGCAGC 362

RESULT 2

AW750670 460 bp mRNA EST 28-APR-2000
LOCUS
DEFINITION CM2-CN0039-110100-069-c01 CN0039 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW750670
VERSION AW750670.1 GI:7665602
KEYWORDS EST.

ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 460)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE
COMMENT 20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-CN0039-
110100-069-c01&t3=2000-01-11&t4=1)
Seq primer: puc 18 forward.
Location/Qualifiers

FEATURES

source

1. 460
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CN0039"
/dev_stage="Adult"
/note="Organ: colon, normal; Vector: puc18; Site: 1; SmaI; Site: 2; SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 114 a 96 c 100 g 150 t
ORIGIN

Query Match 67.1%; Score 169; DB 10; Length 460;
Best Local Similarity 97.2%; Pred. No. 1.6e-41;
Matches 172; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 CAAGAATGTGTGTAGTCTTGTGACAAATCTGAGACATGGCGACTGTACCGCTC 60
|||||
Db 135 CAAGAATGTGTGTAGTCTTGTGACAAATCTGAGACATGGCGACTGTACCGCTC 126
OY 51 AATGCACTGAATCAAGCAGCCAGCTTTCTGCTGACAGACAGTTGAGTGGGCTCTG 120
|||||
Db 125 AATGCACTGAATCAAGCAGCCAGCTTTCTGCTGACAGACAGTTGAGTGGGCTCTG 66
OY 121 GTTGGAGTGTGACATTTGACAGTGTGCTGCCATGTACAAAGTGAACATCAATTAAC 177
|||||
Db 65 GTTGGAGTGTGACATTTGACAGTGTGCTGCCATGTACAAAGTGAACATCAATTAAC 9

RESULT 3

AK008659 2933 bp mRNA HTC 05-JUL-2001
LOCUS
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched
library, clone:221006G11, full insert sequence.
ACCESSION AK008659
VERSION AK008659.1 GI:12842987
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male stomach cDNA to mRNA,
clone: 11b:RIKEN full-length enriched mouse cDNA library
clone: 221006G11.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2933)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Methods in enzymology. 303, 19-44 (1999)

REFERENCE 2 (bases 1 to 2933)
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, N., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome research. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE
COMMENT 20499374
11042159

REFERENCE 3 (bases 1 to 2933)
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

TITLE	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishihara, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiyagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Onozawa, K., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Kozaki, Y., Murakatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
JOURNAL	RIKEN Integrated sequence analysis (RISA) system—334-format sequencing pipeline with 384 multicapillary sequencer
REFERENCE	Genome research. 10 (11), 1757-1771 (2000)
AUTHORS	Genome research. 10 (11), 1757-1771 (2000)
TITLE	4 (bases 1 to 2915)
JOURNAL	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
REFERENCE	Functional annotation of a full-length mouse cDNA collection
AUTHORS	Nature 409, 685-690 (2001)
	5 (bases 1 to 2915)
	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome_research@gscc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

cdna library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cdna was primed with a primer [5' GAGAGAGACAAGAGATCCAGACCTCTTTTTTTTTTTTTTNN 3'], cdna was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cdna went through one round of subtraction to Rot = 20.0%. Second strand cdna was prepared with the primer adapter of sequence¹⁵, GAGAGAGACAGCCCCCGCATTTATTTCTCAGTTATTTAAATTAATTCGCCCCCCC 3'. cdna was cleaved with XhoI and SstI. Cloning sites, 5' end: xhoI, 3' end: SstI. Host: SOLR.

FEATURES	SOURCE
location/Qualifiers	1. .2915
/organism="Mus musculus"	
/strain="C57BL/6J"	
/db_xref="taxon:10090"	
/db_xref="MGI:1346342"	
/db_xref="MGI:1893339"	
/clone="1810012P03"	
/sex="male"	
/tissue_type="pancreas"	
/clone_library="Riken full-length enriched mouse cDNA library"	
/dev_stage="10 day old"	
BASE COUNT	838 a 720 c 695 g 662 t
ORIGIN	

Query Match	Similarity	64.5%	Score	162.6	DB	12	Length	2915	
Best Local	Similarity	78.3%	Pred	No. 2.6e-39					
Matches	195	Conservative	0	Mismatches	54	Indels	0	Gaps	0

QY	1	CAAGATTGTGTGTTTACTCCTTGACAAATGTGGAAGCATGCGGACTGTGTAACCGCCCTC	60
DB	941	CAAGATTTGTGCTTATCTTCTTGATTAAGTCGGGAGCATCTGGAACATCATATCTCTTT	1000

QY	61	AATGACGCTAAATCAAGCAGCGCCAGCTTTTCCGTGCGCAGACAGTTGAGCTGGGGCTCTGG	120
Db	1001	GACCGAATTAATTCAGGCAAGCCGGCTTTTCTCTGCTGCAGACTGTGGACAGGATCTCTGG	1060
QY	121	GTGTGGATGCTGACATTTTGACAGTGTCTGCCATGTACAAAGTGAAGTCAATACATTAAC	180
Db	1061	GTCGGATGCTGATACCTTTTGACAGTGTCTGCTCATGTACAAACGCAATCAAACTTTTAAC	1120
QY	181	AGTGGCAGTGCAGCGGAGACACATCGCCCAAAAGATTACCTGGAGGAGCTTCGAGAGGAGC	240
Db	1121	AGTGGTGGTGCAGCAGAGATCTGCTGATCAACGACATTACCCACAGTATCTGCAGGAGGACA	1180
QY	241	TCCATCTGC	249
Db	1181	TCTATATGC	1189

RESULT	5
EG965023	
LOCUS	704 bp mRNA
DEFINITION	60283166BF1 NOT_CGAP_Co24 Mus musculus cdna clone IMAGE:4986281 mRNA sequence.
ACCESSION	EG965023
VERSION	EG965023.1 GI:14352660
KEYWORDS	EST.
SOURCE	house mouse,
ORGANISM	Mus musculus

REFERENCE	(bases 1 to 704)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

FEATURES

Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MCC clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10995 row: 9 column: 18
High quality sequence stop: 663.

Location/Qualifiers

```

/organism="Mus musculus"
/strain="FVB/N"
/ID_xref="taxon:10090"
/clone="IMAGE:4986281"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: Not
Site_2: SalI; Cloned unidirectionally. Primer: 01lgo
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI-CGAP Library."

```

	Query Match	63.7%	Score 160.6;	DB 11;	Length 704;
	Best Local Similarity	78.18;	Pred. No. 7.2e-39;	Mismatches 193; Conservative	Indels 0; Gaps
CY	3 AAGATTGTGCTTACTCCTTGGCAATCGTAGACATGGCAGTCAGTAACGCCGCCTAA	62			
DB	26 AAGATTGTGCTTACTCTTGTAAGTCGCGAGCATGCTGAACGATGATCGCTTAA	85			
CY	63 TCGACTGATTAACGACGACGCTTTCCCTGCTGCAGACAGTTGAGAECTGGGCTCTGGGT	122			
DB	36 CCGATGATATAGGCAAGCCGGCTTTTTCTCTCTGCACATCTGTGAGCACAGGATCTCTGGT	145			

RESULT	LOCUS	DEFINITION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	BASE COUNT	ORIGIN
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												
647	TATATGC 653												
243	CATCTGC 249												

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:53:10 ; Search time 521.98 Seconds

(without alignments)
109.338 Million cell updates/sec

Title: US-09-049-696-6

Perfect score: 252
Sequence: 1 CAAAGATTGCTGTTTACT.....GAGGACGTCATCTGCACG 252

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11328999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30.8	12.2	1455	1 US-08-832-883-58	Sequence 58, Appl
2	30.8	12.2	1455	2 US-08-832-877-58	Sequence 58, Appl
3	30.4	12.1	2047	4 US-09-345-468-1	Sequence 1, Appl
4	28.2	11.2	3283	4 US-09-061-709-8	Sequence 8, Appl
5	28.2	11.2	3412	4 US-09-061-709-6	Sequence 6, Appl
6	28.2	11.2	3581	2 US-08-738-349-1	Sequence 1, Appl
7	28.2	11.2	40352	3 US-08-846-111D-15	Sequence 15, Appl
8	27.8	11.0	4138	1 US-08-447-411-75	Sequence 75, Appl
9	27.8	11.0	4138	2 US-08-662-227-33	Sequence 33, Appl
10	27.8	11.0	4138	4 US-09-017-947-33	Sequence 33, Appl
11	27.8	11.0	5211	1 US-08-447-411-1	Sequence 1, Appl
12	27.8	11.0	5924	1 US-08-447-411-44	Sequence 44, Appl
13	27.8	11.0	5948	2 US-08-662-227-1	Sequence 1, Appl
14	27.8	11.0	5948	4 US-09-017-947-1	Sequence 1, Appl
15	27.6	11.0	343	4 US-08-905-223-169	Sequence 169, App
16	27.6	11.0	2990	1 US-08-572-951-1	Sequence 1, Appl
17	27.6	11.0	8655	3 US-09-075-272-1	Sequence 1, Appl
18	27.4	10.9	1130	1 US-07-864-004B-1	Sequence 1, Appl
19	27.4	10.9	1130	1 US-08-251-937A-1	Sequence 1, Appl
20	27.4	10.9	1130	1 US-08-212-133A-5	Sequence 5, Appl
21	27.4	10.9	1130	1 US-08-474-503-3	Sequence 3, Appl
22	27.4	10.9	1130	2 US-08-670-707A-3	Sequence 3, Appl
23	27.4	10.9	1130	4 US-09-037-601-3	Sequence 3, Appl
24	27.4	10.9	1130	5 PCT-US93-03275-1	Sequence 1, Appl
25	27.4	10.9	1130	5 PCT-US94-13200-3	Sequence 3, Appl
26	27.4	10.9	1623	1 US-08-121-202-3	Sequence 3, Appl
27	27.4	10.9	4334	2 US-08-670-707A-38	Sequence 38, Appl

28	27.4	10.9	4334	4 US-09-037-601-38	Sequence 38, Appl
29	27.4	10.9	4700	4 US-09-150-460B-9	Sequence 9, Appl
30	27.4	10.9	6402	2 US-08-670-707A-36	Sequence 36, Appl
31	27.4	10.9	6402	4 US-09-037-601-36	Sequence 36, Appl
32	27	10.7	262	1 US-08-222-177A-46	Sequence 46, Appl
33	26.8	10.6	479	4 US-08-927-219-134	Sequence 134, App
34	26.8	10.6	4138	1 US-08-323-474-1	Sequence 1, Appl
35	26.8	10.6	4138	5 PCT-US93-06093-1	Sequence 1, Appl
36	26.8	10.6	4905	1 US-07-978-895-3	Sequence 3, Appl
37	26.8	10.6	4905	1 US-08-473-119-3	Sequence 3, Appl
38	26.8	10.6	4905	2 US-08-475-352-3	Sequence 3, Appl
39	26.8	10.6	4975	4 US-09-630-706-3	Sequence 3, Appl
40	26.8	10.6	6254	4 US-08-927-219-126	Sequence 126, App
41	26.6	10.6	3254	1 US-08-162-809-15	Sequence 15, Appl
42	26.6	10.6	4451	3 US-08-717-294-42	Sequence 42, Appl
43	26.6	10.6	4403765	4 US-09-103-840A-2	Sequence 2, Appl
44	26.6	10.6	4411529	4 US-09-103-840A-1	Sequence 1, Appl
45	26.4	10.5	278	1 US-08-222-177A-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-08-832-883-58
Sequence 58, Application US/08832883
Patent No. 5807681
GENERAL INFORMATION:

APPLICANT: Giordano, Antonio

APPLICANT: Baldi, Alphonso

TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS

NUMBER OF SEQUENCES: 115

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEIDEL, CONDA, LAVORGNA & MONACO, P.C.

STREET: Suite 1800 Two Penn Center Plaza

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19102

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/832,883

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A

REGISTRATION NUMBER: 30,480

REFERENCE/DOCKET NUMBER: 8321-13 US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-8383

TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:

LENGTH: 1455 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-832-883-58

Query Match 12.2% Score 30.8; DB 1; Length 1455;

Best local Similarity 57.1%; Pred. No. 0.55;

Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Cy 38 TTCCTCCGACACACTTACGCTGGGCTCCGCTGGTGGATGCTGACATTTCACACTGCT 147
Db 501 TTCAGGTTTAGCAGACGACATACACCTTAAGTAAGTAGGAGTACCTGTACATTTC 560

OY 148 GCCCATGTACAAAGTGAACATACAGATTAACAGTGG 185
||||||| || ||| ||||| |||||
DB 561 CCCCATGTCAAAAGAGGAAACAGATCATTAATTGTGG 598

RESULT 2

US-08-832-877-58
Sequence 58, Application US/08832877
Patent No. 5840506

GENERAL INFORMATION:

APPLICANT: Giordano, Antonio
TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS OF
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
STREET: Suite 1800 Two Penn Center Plaza
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/832.877
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-13 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-5549
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-832-877-58

Query Match 12.2%; Score 30.8; DB 2; Length 1455;
Best Local Similarity 57.1%; Pred. No. 0.55;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

OY 88 TTCTCTCTGCACAGACGTGAGTGGGTCCTGGTGGGATGTGACATTTCACAGTGGT 147
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 501 TTCAGGGTTTACGACACACATACACCTTAAAGTAGTAGTACCTGTCACATTTC 560
OY 148 GCCCATGTACAAAGTGAACATACAGATTAACAGTGG 185
||||||| || ||| ||||| |||||
DB 561 CCCCATGTCAAAAGAGGAAACAGATCATTAATTGTGG 598

RESULT 3

US-09-345-468-1/C
Sequence 1, Application US/09345468
Patent No. 6245527

GENERAL INFORMATION:

APPLICANT: Busfield, S.
APPLICANT: Villereal, J.
APPLICANT: Jandriot-Perrus, M.
APPLICANT: Vainchenker, W.
TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
FILE REFERENCE: 7853-147
CURRENT APPLICATION NUMBER: US/09/345,468

CURRENT FILING DATE: 1999-06-30

NUMBER OF SEQ ID NOS: 24
SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 1
LENGTH: 2047

TYPE: DNA

ORGANISM: Homo sapiens
US-09-345-468-1

Query Match 12.1%; Score 30.4; DB 4; Length 2047;
Best Local Similarity 54.5%; Pred. No. 0.89;
Matches 61; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

OY 125 GATGTGACATTGTGACGTGCGCCATGTACAAAGTGAACATACAGATTAACAGTGG 184
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1299 GATGGGGTCTCCACAGATTCCTTCATCCCAATGGAGGCGCCCTCAGACAGAGGCG 1240

OY 135 GCAGTGCAGAGGACACACTGCCCAAAAGATTACTGCAGCAGCTTCAGAGG 236
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1239 AGACAGACAGACAGACACTGCGGACGCGGCTCCCTGATGGAACACGAGAGG 1188

RESULT 4

US-09-061-709-8
Sequence 8, Application US/090617098
Patent No. 6297364

GENERAL INFORMATION:

APPLICANT: Chen, Yao-Tseng
APPLICANT: Gure, Ali
APPLICANT: Tseng, Solam
APPLICANT: Stockert, Elisabeth
APPLICANT: Jager, Elke
APPLICANT: Knuth, Alexander
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
FILE REFERENCE: LUD 5538
CURRENT APPLICATION NUMBER: US/09/061,709B
CURRENT FILING DATE: 1998-04-17
NUMBER OF SEQ ID NOS: 8
SEQ ID NO 8
LENGTH: 3283
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-061-709-8

Query Match 11.2%; Score 28.2; DB 4; Length 3283;
Best Local Similarity 59.3%; Pred. No. 6.3;
Matches 48; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 133 ACATTGACAGTGTGCGCCATGTACAAAGTGAACATACAGATTAACAGTGGAC 192
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 391 atatggcagtgtagtaatgtggaacaagtaacacagacaagaacccgctgtgcaa 440
OY 193 AGGACACACTCGCCCAAGA 213
||||| ||||| ||||| ||||| |||||
DB 441 cgtcacatagcaacaagaga 461

RESULT 5

US-09-061-709-6
Sequence 6, Application US/090617098
Patent No. 6297364

GENERAL INFORMATION:

APPLICANT: Chen, Yao-Tseng
APPLICANT: Gure, Ali
APPLICANT: Tseng, Solam
APPLICANT: Stockert, Elisabeth
APPLICANT: Jager, Elke
APPLICANT: Knuth, Alexander

CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/017,947
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/662,227
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 4138 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-017-947-33

Query Match 11.0%; Score 27.8; DB 4; Length 4138;
Best Local Similarity 53.2%; Pred. No. 9.6;
Matches 59; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

OY 102 AGTTGAGCTGGGCTCCTGGTGGATGTCACATTTGACAGTCTGCCCATGTACAAAG 161
DB 755 AGGGATCCAGCTGCTGGATGCTGCTGGTGGATGACAAACAGATATGTTCTCA 814
OY 162 TGAACCTCATACAGATAAAGTGGCAGTGACAGGACACTGCCCAAG 212
DB 815 TGATTAATATTAAGATTAGCCAACTAGATATGGACACATAGAAAAG 865

RESULT 11
US-08-447-411-1
Sequence 1, Application US/08447411
Patent No. 5773243
GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, No. 5773243man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-101-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5211 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 9..4961
US-08-447-411-1

Query Match 11.0%; Score 27.8; DB 1; Length 5211;
Best Local Similarity 53.2%; Pred. No. 11;
Matches 59; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

OY 102 AGTTGAGCTGGGCTCCTGGTGGATGTCACATTTGACAGTCTGCCCATGTACAAAG 161
DB 1730 AGGGATCCAGCTGCTGGATGCTGCTGGTGGATGACAAACAGATATGTTCTCA 1789
OY 162 TGAACCTCATACAGATAAAGTGGCAGTGACAGGACACTGCCCAAG 212
DB 1790 TGATTAATATTAAGATTAGCCAACTAGATATGGACACATAGAAAAG 1840

RESULT 12
US-08-447-411-44
Sequence 44, Application US/08447411
Patent No. 5773243
GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5773243man F.

RESULT 13
US-08-662-227-1.
: Sequence 1, Application US/08662227
: Patent No. 5922320
: GENERAL INFORMATION:
: APPLICANT: VOGEL, CARL-WITHELM
: APPLICANT: BREDEHORST, REINHORST
: APPLICANT: KOCK, MICHAEL
: APPLICANT: FRITZINGER, DAVID
: TITLE OF INVENTION: RECOMBINANT PROCVF
: NUMBER OF SEQUENCES: 39
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESS: P.C.
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/662,227
: FILING DATE: 14-JUN-1996
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: OBION, NORMAN F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 1126-0107-0X
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-413-3000
: TELEFAX: 703-413-2220

```
Query: Match : 11.0%; Score 27.6; DB 4; Length 5948;
: Best Local Similarity 53.2%; Pred. No. 11;
```


Matches 59: Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 102 ACTGAGCTGGGCTCGGTGGATGCTGACATTTGACACTGCTGCCATGTACAAG 161

DB 1707 AGGGATGCAGCTGCTGGCTTGGCTTGTGGCTGTGACAAAGCATATATCTCTCA 1766

QY 162 TGAAGTCTATACGATTAACAGTGGACAGGACAGGACATCGCCAAAG 212

DB 1767 TGATTAATATAGATTAGCCAGCTAAGATATGGACACAAATAGAAAGAG 1817

RESULT 15

US-08-905-223-169/c

Sequence 169, Application US/08905223

Patent No. 6222029

GENERAL INFORMATION:

APPLICANT: Edwards, Jean-Baptiste D.

APPLICANT: Duclert, Aymeric

APPLICANT: Lacroix, Bruno

TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS

NUMBER OF SEQUENCES: 503

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Win95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/905,223

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Israel, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 169:

SEQUENCE CHARACTERISTICS:

LENGTH: 343 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: DOUBLE

TOPOLOGY: LINEAR

MOLECULE TYPE: CDNA

ORIGINAL SOURCE:

ORGANISM: Homo Sapiens

TISSUE TYPE: Brain

FEATURE:

NAME/KEY: other

LOCATION: 104..336

IDENTIFICATION METHOD: blastn

OTHER INFORMATION: identity 96

OTHER INFORMATION: region 1..233

OTHER INFORMATION: id H07998

OTHER INFORMATION: est

FEATURE:

NAME/KEY: other

LOCATION: 110..336

IDENTIFICATION METHOD: blastn

OTHER INFORMATION: identity 96

OTHER INFORMATION: region 1..227

OTHER INFORMATION: id W37530

OTHER INFORMATION: est

FEATURE:

NAME/KEY: other

LOCATION: 110..336

IDENTIFICATION METHOD: blastn

OTHER INFORMATION: identity 96

OTHER INFORMATION: region 1..227

OTHER INFORMATION: id R79812

OTHER INFORMATION: est

FEATURE:

NAME/KEY: other

LOCATION: 110..336

IDENTIFICATION METHOD: blastn

OTHER INFORMATION: identity 96

OTHER INFORMATION: region 1..227

OTHER INFORMATION: id N24900

OTHER INFORMATION: est

FEATURE:

NAME/KEY: other

LOCATION: 110..336

IDENTIFICATION METHOD: blastn

OTHER INFORMATION: identity 96

OTHER INFORMATION: region 1..227

OTHER INFORMATION: id R34849

OTHER INFORMATION: est

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 65..112

IDENTIFICATION METHOD: Von Heljne matrix

OTHER INFORMATION: score 12.5

OTHER INFORMATION: seq FVLLALVAVGLG/NE

US-08-905-223-169.

Query Match

Best Local Similarity

Matches 60: Conservative

11.0%; Score 27.6; DB 4; Length 343;

52.6%; Pred. No. 3.4;

Mismatches 54; Indels 0; Gaps 0;

QY 108 TGGTAACCGCCCTCATCGACTGATGATGACAGGAGCCAGCTTTCTGCTGACAGACTTGA 107

DB 317 TGTAACTCTCCCTTACACATGACATGACGCTAGCCGAGAGAGATGAACAGGTTACCA 258

QY 108 GCTGGGGTCCCTGGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 161

DB 257 CTGCGAGTCTCTGCGCAAGAAAGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 204

Search completed: April 3, 2002, 20:53:31

Job time: 41769 sec

This Page Blank (uspto)

XX	Nakanishi A, Morita S:
FI	
XX	WPI; 2001-355935/37.
DR	P-PSDB; AAB73716.
DR	
XX	New antisense nucleotide, useful for treatment and prevention of
PT	bronchial asthma and chronic obstructive pulmonary disease -
PT	
XX	Claim 3; Page 82-84; 104pp; Japanese.
PS	
XX	The invention relates to an antisense nucleotide targeted to the mouse
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC	the C1CAl gene (coding sequence shown in AAH46102). The invention also
CC	relates to an antibody specific for the Gob-5 protein, medical and
CC	diagnostic compositions containing the antisense nucleotide or the
CC	antibody, and methods and kits for screening for compounds which inhibit
CC	the protein. Gob-5 and C1CAl are proteins expressed by goblet cells.
CC	The antisense oligonucleotides and antibody are therefore useful for the
CC	treatment and prevention of bronchial asthma and chronic obstructive
CC	pulmonary disease. The present sequence represents the human C1CAl
CC	gene coding sequence.
SQ	
Sequence	2742 BP; 833 A; 616 C; 623 G; 670 T; 0 other;
Query Match	100.0%; Score 252; DB 22; Length 2742;
Best Local Similarity	100.0%; Pred. No. 8,6e-73;
Matches	252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 CAAGAATTGTGTTTACTCTTCACAAATCTGAGAGCATGGCGACTGTGAACCGCTC 60
Db	 caaaagaattgttcttagtccttgacaacatctggaagatgycgactytaaacycctcc 969
OY	61 AATGCATGTAATCAACGACAGCGACGCTTTCCGCTGCACACAGTTGAGCTGGGGTCCCTGG 120
Db	 aatgcacgtaaccaagccaagccaccttcccgcgcgacacagttgagctggggtccctgg 1029
OY	121 GTTGCGATGCTGACATTTTGACAGTAGTGCCTGCCCATGTATACAAGTGAATCTATACAGATAAAC 180
Db	 gttcgcatgtgtgatactttgacagtctgccccatgtaacaagyaaactcatacagattaaac 1089
OY	181 AGTGGCAGTGCACAGAGGACACACTCGCCAAAATAATTCCTGCGACGACTTCAGAGAGGGAGC 240
Db	 agtggcagtggaacaggcacactcgccaaaataattaccctgcagcagcttcacaggaagggacg 1149
OY	241 TCCATCTCGACG 252 tccatctcgacg 1161
Db	1150 tccatctcgacg 1161
RESULT	2
AAF81927	
ID	AAF81927 standard; cDNA; 2745 BP.
XX	
AC	AAF81927;
XX	
DT	13-JUN-2001 (first entry)
XX	
DE	Human ICACC-1 nucleotide sequence.
XX	
KM	ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
KM	interleukin 9 induced calcium activated chloride channel; IL-9;
KM	calcium activated chloride channel; anti-allergic; anti-asthmatic;
KM	anti-inflammatory; immunomodulatory; cystic fibrosis;
KM	inflammatory bowel disease; autoimmune disease; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FH	Location/Qualifiers
FT	CDS 1..2745
FT	/*tag= a
FT	/product= "ICACC-1"

FT	XX		/note= "IL-9 induced calcium activated chloride channel"
PN	XX	W09944620-A1.	
PD	XX	10-SEP-1999.	
PF	XX	03-MAR-1999; 99WO-US04703.	
PR	XX	03-MAR-1998; 98US-0076815.	
PA	XX	(MAGA-) MAGALIN PHARM INC.	
PI	XX	Holroyd RJ, Levitt RC, Maloy WL, Louhed J, McLane M;	
PI	XX	Nicolaides NC, Zhou Y, Dong Q;	
DR	XX	WPI; 1999-550979/46.	
DR	XX	P-PSDB; AAB74824.	
PT	XX	New nucleic acid encoding calcium activated chloride channel, used to	
PT	XX	identify, e.g. specific modulators for treating atopic allergy -	
PS	XX	Claim 1; Fig 4B; 75pp; English.	
CC	XX	The present sequence encodes the human interleukin 9 (IL-9) induced	
CC	XX	calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins	
CC	XX	have anti-allergic, anti-asthmatic, anti-inflammatory and	
CC	XX	immunomodulatory activities. Compounds (A) that downregulate ICACC are	
CC	XX	used to alleviate asthma (or more generally atopic allergy), while those	
CC	XX	(B) that upregulate ICACC are used to treat cystic fibrosis (CF) and	
CC	XX	inflammatory bowel disease (IBD) (or other autoimmune diseases).	
CC	XX	Measuring levels of ICACC-1 can be used in the diagnosis of asthma	
CC	XX	(increased levels) or IBD (reduced levels), also for monitoring	
CC	XX	treatment of these conditions. The ICACC proteins can be used:	
CC	XX	(1) to raise specific antibodies (Ab); useful: (a) as immunoassay	
CC	XX	reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as	
CC	XX	(A) to disrupt bonding between ICACC and its ligands; and (iii) to	
CC	XX	identify modulators and binding partners. ICACC polynucleotides can be	
CC	XX	used to generate transgenic animals or recombinant cells, used to screen	
CC	XX	for antagonists, also as a source of therapeutic antisense agents or	
CC	XX	diagnostic probes (for quantifying mRNA expression, e.g. for	
CC	XX	identification of modulators).	
SQ		Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other:	
		Query Match 100.0%; Score 252; DB 20; Length 2745;	
		Best Local Similarity 100.0%; Pred. No. 8.ee-73;	
		Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	CANAAGATGTGTTTGTAGTCCTTACCAAAATTGGAACATGGCGAGTGTAACGCCCTC 60	
DB	910	caaaagaattgttgtttagtccttgacaatctggaagcaatggcgactgtaaccgcctc 969	
QY	'61	AATGCAGTAATCAAGCAGCGCCAGCTTTTCCGTGCTGCAGACAGTTGAGCTGGGGCTCTGG 120	
DB	970	aatgcacgaataccaagccaggccagcttttccctgcgcgaagtgtagctggggtccgcg 1029	
QY	121	GTTGGATGCTGACATTTGACAGTCTGCCCATGTGACAAAGTGAACCTATACAGATAAAC 180	
DB	1030	gttggatgctgacatttgacatttgacagtctgcgccatgfacaaaagactcatacaagtaaac 1089	
QY	181	AATGGCAGTAGCAGGAGACACACTGCGCAAAGAATTAACCTGCGACGAGCTTCAGGAGGAGC 240	
DB	1090	aatggcagtagcagaaggaacacacctgcgcaaaaagattaccctgcagcagcttcaaggagacg 1149	
QY	241	TCCATCTGCAGC 252	
DB	1150	tccatctgcagc 1161	
RESULT	3		
ID	AAH46124		
	AAH46124 standard; CDNA: 2825 BP.		

XX	AAH46124;
AC	
XX	
DT	11-SEP-2001 (first entry)
XX	
DE	Human CLCA1 cDNA, SEQ ID NO:26.
XX	
KW	Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening; expression inhibition; antisense therapy; gene therapy;
KW	chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
XX	ss.
OS	Homo sapiens.
XX	
EH	Key
FT	Location/Qualifiers
FT	23..2767
FT	/tag= a
FT	/product= "Human CLCA1"
FT	/transl_except= (pos:476..478, aa:Lys)
XX	
PN	WO200138530-A1.
XX	
PD	31-MAY-2001.
XX	
PE	22-NOV-2000; 2000WO-JP08232.
XX	
PR	24-NOV-1999; 99JP-033479.
XX	27-APR-2000; 2000JP-0127589.
PA	(TAKE) TAKEDA CHEM IND LTD.
PI	Nakanishi A, Morita S;
DR	WPI: 2001-355935/37.
DR	P-PSDB: AAB73716.
XX	
PT	New antisense nucleotide, useful for treatment and prevention of
PT	bronchial asthma and chronic obstructive pulmonary disease
XX	
PS	Example 5; Page 92-94; 104pp; Japanese.
XX	
CC	The invention relates to an antisense nucleotide targetted to the mouse
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC	the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC	relates to an antibody specific for the Gob-5 protein, medical and
CC	diagnostic compositions containing the antisense nucleotide or the
CC	antibody, and methods and kits for screening for compounds which inhibit
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC	The antisense oligonucleotides and antibody are therefore useful for the
CC	treatment and prevention of bronchial asthma and chronic obstructive
CC	pulmonary disease. The present sequence represents human CLCA1 cDNA.
XX	
XX	
Sequence	2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;
Query Match	100.0%; Score 252; DB 22; Length 2825;
Best Local Similarity	100.0%; Pred. No. 8.7e-73;
Matches 252; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 CAAGAAATTTGCTGTTTAACTCTCTTGACAAATCTGGAAGCATGGCCACTGTGTAACCGCCCTC 60
DB	
OY	932 caaagaatgtgtgttgcctctgtcaacaactctggaagcattgacgactgtgtacacgcctc 991
OY	61 AATCATGTAATCAACAGCAGGCGCAGCTTTTCTGCTGCACAGCATGGTGGGGTCTCTGG 120
DB	
OY	992 aatgcactgcaatcaagcagcgccagctttccctgcctgcagacacggttgagctggggtccctgg 1051
OY	121 GTTGGGATGTGACATTTTGACAGTGTCTGCGCCATGTACAAAGTGAACCTCATACAGATAAC 180
DB	
OY	1052 gtctggatgtgtacatttcagtcagctgtcgccactgtacaagaagtgaaactcatacagataaac 1111
OY	181 AATTGCAGTGAAGGACACACTCCCAAAATTAACCTGCAGCAGCTTCAGAGGGAGG 240

```
Db      1112 agcgccagctgacagggaaccactgcgccaaagaattacctcggcgcagtccatcgaggagcaag 1171
OY      281 TCACATCTGCAGC 252
        |||||
Dd      1162 ttcacatctgcagc 1183
RESULT: 4
ID      AAH34879 standard; cDNA; 2854 BP.
XX      AC   AAh34879;
KW      Homo sapiens;
OS      Homo sapiens;
PN      MQ200122920-A2.
PF      28-SEP-2000; 2000WO-US26524.
PR      29-SEP-1999; 99US-0157137.
RR      03-NOV-1999; 99US-0163280.
PA      (HJMA-) HUMAN GENOME SCI INC.
PI      Ruben SM, Barash SC, Birse CE, Rosen CA;
DR      WRL: 2001-235357/24.
RSDB: AAG75474.
Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
useful for preventing, diagnosing and/or treating colorectal cancers -
Chaim I; Page 3462-3463; 9803pp; English.
AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
cancer-associated nucleic acid molecules (N) and proteins (P), where
the proteins are collectively known as colon cancer antigens. The colon
cancer antigens have cytostatic activity and can be used in gene
therapy and vaccine production. N and P may be used in the prevention,
diagnosis and treatment of diseases associated with inappropriate P
expression. For example, N and P may be used to treat disorders
associated with decreased expression by rectifying mutations or deletions
in a patient's genome that affect the activity of P by expressing P.
Inactivating proteins or to supplement the patients own production of P.
Additionally, N may be used to produce the colon cancer-associated PS,
by inserting the nucleic acids into a host cell and culturing the cell
to express the proteins. N and P can be used in the prevention, diagnosis
and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
and AAB77789 represent sequences used in the exemplification of the
present invention.
N.B. Pages 666 to 682 and page 7053 of the sequence listing were
mislabelled at time of publication, meaning no sequences are present for
SEQ ID NO:1027 to 1052, 7921 and 7922.
Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;
```


PR 29-JUN-2000; 2000US-0609448.
PR 28-AUG-2000; 2000US-0649811.
PA (CORI-) CORIXA CORP.
PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX WPI: 2001-441847/47.
PT Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer -
XX
XX
PS Claim 2; Page 425-426; 472pp; English.
XX The present invention describes colon tumour associated proteins (I) and
XX the polynucleotides (II) that encode them. (I) have cytostatic activity.
XX CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
XX CC (II) may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate colon tumour associated protein (TCAP)
XX CC expression, such as colonic cancer. For example, (I) and (II) may be
XX CC used to treat disorders associated with decreased expression by
XX CC rectifying mutations or deletions in a patient's genome that affect the
XX CC activity of TCAPs by expressing inactive proteins or to supplement the
XX CC patients own production of them. Additionally, (II) may be used to
XX CC produce the TCAP proteins, by inserting the nucleic acids into a host
XX CC cell culturing the cell to express the protein. (II) and its
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC polymerase chain reaction (PCR) and hybridisation assays to detect and
XX CC quantitate the presence of similar nucleic acids in samples, and
XX CC also be used as antigens in the need of restorative therapy. (I) may
XX CC be used as antigens in the production of antibodies against TCAPs
XX CC and in assays to identify modulators of TCAP expression and activity.
XX CC Anti-(I) antibodies and antagonists may also be used to down regulate
XX CC TCAP expression and activity. The anti (I) antibodies may also be used
XX CC as diagnostic agents for detecting the presence of TCAPs in samples
XX CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
XX CC and AA24494 to AA24533 represent nucleotide and amino acid sequences
XX CC given in the exemplification of the present invention.
XX
XX Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;
SO
Query Match 100.0%; Score 252; DB 22; Length 3311;
Best Local Similarity 100.0%; Pred. No. 9.3e-73;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CAAAGCAATTGTGTGTTAGTCTTGAACAATCTGGAAGCATGGCCACTGTTACCGCTC 60
DB 1261 caaagaattgtgttcttgccttgcacaatctggaagcatggtgactgtaaccgctc 1320
OY 61 AATGAGTGAATCAAGCAGGCGAGCTTTCTCGTGCAGACAGTGTGAGTGGGCTCTGG 120
DB 1321 aatgagtgaaatcaagcagcgagctttctcgtcgacagagcttgagctgggtctcgg 1380
OY 121 GTTGGATGCTGACATTTGACAGTCTGCCATGTACAAAAGTGAATCATACAGATAAAC 180
DB 1381 gttggatgctgacatttgacagtgctgcccatgtacaagaagtaacatacagataaac 1440
OY 181 AGTGCAGTGCAGAGGACACACTCGCCAAAAGATTACCTGCAGACAGCTTCAGAGGAGC 240
DB 1441 agtgcagtgacagagacactcgcccaaaagattacctgcagagcttcagagagagac 1500
OY 241 TCATCTGCAGC 252
DB 1501 tcatctgcagc 1512
RESULT 9
AAH33285
ID AAH33285 standard; cDNA; 2867 BP.
XX
AC AAH33285;

XX
DE 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:341.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 1; ss.
XX
CS Homo sapiens.
XX
PN W0200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI: 2001-235357/24.
XX P-PSDB: AAG73854.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX
XX Claim 1; Page 2452-2453; 9803pp; English.
XX
XX AA132943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing
XX inactive proteins or to supplement the patients own production of P.
XX Additionally, N may be used to produce the colon cancer-associated P,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAB77789 represent sequences used in the exemplification of the
XX present invention.
XX
XX N.3. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX
XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;
SO
Query Match 95.6%; Score 241; DB 22; Length 2867;
Best Local Similarity 99.6%; Pred. No. 3.6e-65;
Matches 252; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
OY 1 CAAAGCAATTGTGTGTTAGTCTTGAACAATCTGGAAGCATGGCCACTGTTACCGCTC 60
DB 947 caaagaattgtgttcttgccttgcacaatctggaagcatggtgactgtaaccgctc 1006
OY 51 AATGAGTGAATCAAGCAGGCGAGCTTTCTCGTGCAGACAGTGTGAGTGGGCTCTCG 120
DB 1007 aatgagtgaaatcaagcagcgagctttctcgtcgacagagcttgagctgggtctcgg 1066
OY 121 GTTGGATGCTGACATTTGACAGTCTGCCATGTACAAAAGTGAATCATACAGATAAAC 180
DB 1067 gttggatgctgacatttgacagtgctgcccatgtacaagaagtaacatacagataaac 1126
OY 181 AGT-GGCAGTGCAGAGGACACACTCGCCAAAAGATTACCTGCAGACAGCTTCAGAGGAGC 239
DB 1127 agtgcagtgacagagacactcgcccaaaagattacctgcagagcttcagagagagac 1186

OY 240 GTCATCTGCAGC 252
|||||
Db 1187 gtccatctgcagc 1199

RESULT 10

AAH46101
ID AAH46101 standard; DNA: 2739 BP.

AC AAH46101;

DT 11-SEP-2001 (first entry)

DE Mouse Gob-5 coding sequence, SEQ ID NO:3.

KW Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
KW expression inhibition; antisense therapy; gene therapy; bronchial asthma;
KW chronic obstructive pulmonary disease; antiasthmatic; ds.

OS Mus sp.

FT Key Location/Qualifiers
FT CDS 1..2739
FT /*tag= a
FT /product= "Mouse Gob-5"

FT /note= "No stop codon given in the specification"

XX WO200138530-A1.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-JP08232.

XX 24-NOV-1999; 99JP-0333479.
XX 27-APR-2000; 2000JP-0127589.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Nakanishi A, Morita S;

XX WPI: 2001-355935/37.
XX P-PSDB: AAB73715.

XX New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease -

XX Claim 3; Page 80-82; 104pp; Japanese.

XX The invention relates to an antisense nucleotide targeted to the mouse
XX Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX relates to an antibody specific for the Gob-5 protein, medical and
XX diagnostic compositions containing the antisense nucleotide or the
XX antibody, and methods and kits for screening for compounds which inhibit
XX the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX The antisense oligonucleotides and antibody are therefore useful for the
XX treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents the mouse Gob-5
XX gene coding sequence.

XX Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other;

Query Match 65.2%; Score 164.2; DB 22; Length 2739;
Best Local Similarity 78.7%; Pred. No. 6.1e-44;

Matches 196; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

OY 1 CAAGAATTGTGTGTTGCTTGCACAAATCTGAGAGATGGCGTGTAAACGGCCCTC 60
|||||
Db 913 caagaattgtgtgcttgcctgataagtcgagcagtcgacgagtcgctt 972

OY 61 AATGCACTGATCAAGCAGCCAGCTTTTCTGCTGCAGACAGTTAGCTGGGCTCTCG 120
|||
Db 983 aaccgaatgaatcaagcaagcgccttcttcctgcctgcagcagtcgagaggtactcgg 1032
|||
OY 1131 GTTGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
|||
Db 1033 gtccgagatgagccttcagcagtcgctgctcctatgataaagcaactcaaaagttaaac 1092
|||
OY 147 AGTGCAGTGCAGCAGGAGCAGCTCCGCAAAAGATTACCTGACAGCTTCAGAGGAGC 240
|||
Db 1093 agtgcagtcgacagagatctgtgatcaagcacttaaccagatctgcagagagaca 1152
|||
OY 241 TCCATCTGC 249
|||
Db 1153 tctatatgc 1161

RESULT 11

AAH46120
ID AAH46120 standard; cDNA: 2843 BP.

AC AAH46120;

DE 11-SEP-2001 (first entry)

DE Mouse Gob-5 cDNA, SEQ ID NO:22.

KW Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
KW expression inhibition; antisense therapy; gene therapy; bronchial asthma;
KW chronic obstructive pulmonary disease; antiasthmatic; ss.

OS Mus sp.

FT Key Location/Qualifiers
FT CDS 15..2756
FT /*tag= a
FT /product= "Mouse Gob-5"

XX WO200138530-A1.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-JP08232.

XX 24-NOV-1999; 99JP-0333479.
XX 27-APR-2000; 2000JP-0127589.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Nakanishi A, Morita S;

XX WPI: 2001-355935/37.
XX P-PSDB: AAB73715.

XX New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease -

XX Example 1; Page 89-91; 104pp; Japanese.

XX The invention relates to an antisense nucleotide targeted to the mouse
XX Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX relates to an antibody specific for the Gob-5 protein, medical and
XX diagnostic compositions containing the antisense nucleotide or the
XX antibody, and methods and kits for screening for compounds which inhibit
XX the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX The antisense oligonucleotides and antibody are therefore useful for the
XX treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents mouse Gob-5 cDNA.

XX Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;

Query Match	65.2%	Score 164.2	DB 22	Length 2843
Best Local Similarity	78.7%	Pred. No. 6.2e-44		
Matches 196	Conservative	0	Mismatches 53	Indels 0
				Caps
1	CAAAGAAATTTGTGTTAGTCTTGTGACAAATCTGGAACATGGACGACTGTGAACCGCTC	60		
Db	927 caaagatctgtgtgacttgcattgtcttgataagtcgccggagatgctgaacgatgatcgtctt	986		
QY	61 AATCGACTGAATCAAGCAGCGCCAGCTTTTCTCTCTCAACACAGTTAGAGTGGGGTCCG	120		
Db	987 aaccgaatgaatcaggaacgcgcgtcttcctcgtgcagactgtgagcaaggatcccg	1046		
QY	121 GTTGCGATGTTGGACATTGTGACAGTGTGCTGCCATGTGCAAAAGTGAACTCATACAGATAAC	180		
Db	1047 gtccggtatgtgtgaccttgcacgtgtcgtccctatgtacaagaagcgaactcaaacgattaac	1106		
QY	181 AGTGGCAGTGCACAGGACACACTGCCCAAAAGATTACTGTCCAGCACTTCGAGAGGGACG	240		
Db	1107 agtgtgtcgtcagagatctcgtatcaagaacttaccacagatcttcgaaggaggaca	1166		
QY	241 TCCATCTGC 249			
Db	1167 tctatagc 1175			
RESULT 12				
ID	AAF81925			
	AAF81925 standard; cDNA; 2931 BP.			
XX	AAF81925:			
AC				
XX				
XX	13-JUN-2001 (first entry)			
DT				
XX				
XX				
DE				
XX	Murine ICAC-1 nucleotide sequence.			
XX				
XX	ICAC-1; ICAC-2; asthma; atopic allergy; asthma-related disorder;			
KM	Interleukin 9 induced calcium activated chloride channel; IL-9;			
KM	calcium activated chloride channel; anti-allergic; anti-asthmatic;			
KW	anti-inflammatory; immunomodulatory; cystic fibrosis;			
KW	inflammatory bowel disease; autoimmune disease; ss.			
OS	Mus sp.			
XX				
XX				
FH	Location/Qualifiers			
FT	8..2749			
FT	CDS			
FT	/*tag= a			
FT	/product= "ICAC-1"			
FT	/note= "IL-9 induced calcium activated chloride channel"			
XX				
PN	W09944620-AL.			
XX				
XX				
PD	10-SEP-1999.			
XX				
PF	03-MAR-1999; 99WO-US04703.			
XX				
PR	03-MAR-1998; 98US-0076815.			
XX				
XX	(MAGN-) MAGNIN PHARM INC.			
PA				
XX	Holroyd KJ, Levitt RC, Maloy WL, Louhed J, McLane M;			
PI	Nicolaides NC, Zhou Y, Dong Q;			
XX				
XX	WPI: 1999-550979/46.			
DR	P-PDB; AAB74822.			
XX				
XX	New nucleic acid encoding calcium activated chloride channel, used to			
PT	identify, e.g. specific modulators for treating atopic allergy			
XX				
XX	Claim 2; Fig 2; 75pp; English.			
PS				
XX				
CC	The present sequence encodes the murine interleukin 9 (IL-9) induced			
CC	calcium activated chloride channel 1 (ICAC-1) protein. ICAC proteins			
CC	have anti-allergic, anti-asthmatic, anti-inflammatory and			

immunomodulatory activities. Compounds (A) that downregulate ICAC are used to alleviate asthma (or more generally atopic allergy), while those (B) that upregulate ICAC are used to treat cystic fibrosis (CF) and inflammatory bowel disease (IBD) (or other autoimmune diseases). Measuring levels of ICAC-1 can be used in the diagnosis of asthma (increased levels) or IBD (reduced levels), also for monitoring treatment of these conditions. The ICAC proteins can be used:

- (i) to raise specific antibodies (Ab);
- (ii) as fragments or mutants; as reagents, and (b) as therapeutic (Ab);
- (iii) as ligands; and (iii) to (C) to disrupt bonding between ICAC and its ligands; and (iii) to identify modulators and binding partners.

ICAC polynucleotides can be used to generate transgenic animals or recombinant cells, used to screen for antagonists, also as a source of therapeutic antisense agents or diagnostic probes (for quantifying mRNA expression, e.g. for identification of modulators).

Sequence 2931 BP; 867 A; 719 C; 684 G; 661 T; 0 other;

Query Match
Best Local Similarity 78.7%; Pred. No. 6.3e-44;
Matches 196; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

..1 CAAGAATTGTGTTTAACTCTTACAAATCTGGAAAGCATGGCGACTGGTAACCCGCTC 60
.....
920 caaaagaattgttgcttgccttaagtcgagggaacgatcagatgatcgtcctt 979
.....
Db 61 AACGCATCAATCAACGACGCCAGCTTTTCCTGCTGCAGACAGTTGAAGTGGGGCTCGG 120
.....
Cy 121 GTTGGATGTGTCATTTTACAGCTGCTGCCATGTACAAATGAACTCATACAGATTAAC 180
.....
Db 1040 gtccggatgttgaccttgacagtcgtcgctcatgacaagaagcaactccaacgcttaaac 1099
.....
Cy 191 AGTGGCAGTAGCACGGGACACACTCGCCAAAAGATTACCCTGCAGCAGCTTCAGAGGAGGACG 240
.....
Db 1100 agtggtgctgacagagatctgctgatacaagcacttacaccagatlatctgcagaggggaca 1159
.....
Cy 241 TCACATCC 249
.....
Db 1160 tctatatcgc 1168
.....

RESULT 13
AA64335
ID AA64335 standard; DNA; 2616 BP.
XX
NC AA64335;
XX
DT 20-DEC-2000 (first entry)
XX
DE Clone 2516888 of a novel gene associated with colon disease.
XX
EM Colon disease; colon cancer gene; carbonic anhydrase; galelectin;
KM Carcinoembryonic antigen; colorectal carcinoma tumour associated antigen;
KM fatty acid binding protein; glutathione peroxidase; guanylin;
KW cyokeratin; cadherin; intestinal mucin; colon cancer; Crohn's disease;
KW metastatic colon cancer; atrophic gastritis; cholecystitis;
KW Irritable bowel syndrome; ulcerative colitis; ss.
XX
OS Helio saplens.
XX
PN W0200050588-A2.
XX
PD 31-AUG-2000.
XX
PF 01-FEB-2000; 2000MO-US02595.
XX
PR 22-FEB-1999; 9905-0255381.
XX
PA (INCY-) INCYTE PHARM INC.
XX

PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 03-MAR-2000; 2000US-0187202.
 PR 25-APR-2000; 2000US-0193997.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 05-JUN-2000; 2000US-0209832.

PA (GETH) GENENTECH INC.

PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ,
 PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;

DR WPI; 2001-183260/18.
 DR P-PSDB; AAB87560.

XX Eighty four nucleic acids encoding PRO polypeptides, useful in
 PT molecular biology, including use as hybridization probes, and in
 PT chromosome and gene mapping.

PS Claim 2; Fig 69; 278pp; English.

XX The present sequence is the coding sequence for a human PRO polypeptide
 CC (secreted and transmembrane). The PRO protein, and PRO agonists, PRO
 CC antagonists or anti-PRO antibodies are useful for preparation of a
 CC medicament useful in the treatment of a condition which is responsive to
 CC the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
 CC protein may also be employed as molecular weight markers for protein
 CC electrophoresis. The PRO coding sequence has applications in molecular
 CC biology, including use as hybridisation probes, and in chromosome and
 CC gene mapping.

SQ Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;

Query Match 55.6%; Score 140.2; DB 22; Length 3265;
 Best Local Similarity 72.7%; Pred. NO. 5.1e-36;
 Matches 181; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

OY 1 CAAGCAATTGTTGTTAGTCTTGTACAAATCTGGAGCATGCGAGCTGTACCGCCCTC 60
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 934 caaagaattggtgcttagtctctgataagtcctggaagcaatggtgtaagacgcgctca 993
 OY 61 AATGCACTGAATCAAGCAGCGCCAGCTTTTCTGCTGACAGACAGTGTAGCTGGGCTCTGG 120
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 994 aatcgatgaatcaagaacaaacatttcctcgtcgacgactgtgaaatgatcctg 1053
 OY 121 GTTGGAGTGTGACATTGTGACAGTCTGCTGCCATGTACAAAGTGAATCTCATACAGATTAAC 180
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1054 gtggggaatggtcacttgatgactgacctatgttaataaagcaataaataaa 1113
 OY 181 AGTGCACTGACAGGAGACACTCGCAAAAGATTACCTGCGAGCTTCAGAGAGGAGC 240
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1114 agcaatgatgaagaacacacactcatgacgagattaccatacctcctggaagaact 1173
 OY 241 TCCATCTGC 249
 Db ||||||||||
 Db 1174 tccatctgc 1182

Search completed: April 4, 2002, 03:17:35
 Job time: 60207 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 02:58:21 ; Search time 22700.8 Seconds

(without alignments)
183.134 Million cell updates/sec

Title: US-09-049-696-6

Perfect score: 252
Sequence: 1 CAAGAATTGTGTGTACT.....GAGGACGTCATCTGCACC 252

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_ov:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result #	No.	Score	Match	Length	DB	ID	Description
1	252	100.0	2022	9	AK024970	AK024970 Homo sapi	
2	252	100.0	2826	9	AF127036	AF127036 Homo sapi	
3	252	100.0	3311	6	AX193489	AX193489 Sequence	
4	252	100.0	3311	9	AF039400	AF039400 Homo sapi	
5	209.2	83.0	35278	9	AF039401	AF039401 Homo sapi	
6	209.2	83.0	164891	2	AL358950	AL358950 Homo sapi	
7	207.6	82.4	113764	9	HS651E10	AL122002 Human DNA	
8	175.4	69.6	3079	4	AF095584	AF095584 Sus scrofa	
9	164.2	65.2	2937	10	AB017156	AB017156 Mus muscu	
10	140.2	55.6	3204	9	AF127035	AF127035 Homo sapi	
11	140.2	55.6	3221	9	AK000072	AK000072 Homo sapi	
12	140.2	55.6	3265	6	AX092338	AX092338 Sequence	
13	109	43.3	3022	10	AF047838	AF047838 Homo sapi	
14	109	43.3	3471	10	AF052746	AF052746 Mus muscu	
15	108.6	43.1	2984	4	BT036445	BT036445 Bos taurus	
16	107.4	42.6	2765	10	AF115852	AF115852 Mus muscu	
17	107.4	42.6	3058	10	AF108501	AF108501 Mus muscu	
18	107.4	42.6	3137	10	BC008147	BC008147 Mus muscu	
19	105.8	42.0	140718	2	AL356270	AL356270 Homo sapi	
20	105.4	41.8	3415	9	AF043976	AF043976 Homo sapi	
21	103.8	41.2	2820	4	AF001263	AF001263 Bos taurus	
22	103.8	41.2	3288	4	AF001262	AF001262 Bos taurus	
23	103.8	41.2	3317	4	AF001261	AF001261 Bos taurus	
24	80.6	32.0	175591	2	AC068071	AC068071 Homo sapi	
25	53.8	21.3	2832	9	AF127980	AF127980 Homo sapi	
26	53.8	21.3	2970	9	AF043977	AF043977 Homo sapi	
27	53.8	21.3	3604	9	AB026833	AB026833 Homo sapi	
28	53.8	21.3	4077	6	AX054697	AX054697 Sequence	
29	37.4	14.8	175378	2	AL133323	AL133323 Homo sapi	
30	36	14.3	179481	2	AC034100	AC034100 Mus muscu	
31	35.8	14.2	40442	1	SC18	AL132644 Streptomy	
32	35.4	14.0	2007	6	AX073569	AX073569 Sequence	
33	34.8	13.8	40167	8	AP002063	AP002063 Arabidops	
34	34.8	13.8	85657	8	AC069473	AC069473 Arabidops	
35	34.8	13.8	112099	9	AL162727	AL162727 Human DNA	
36	34.4	13.7	113872	9	HS5769	Z95116 Human DNA s	
37	34.2	13.6	144818	2	AL139343	AL139343 Human DNA	
38	34.2	13.6	162486	2	AC044779	AC044779 Homo sapi	
39	34	13.5	181481	2	AP003038	AP003038 Homo sapi	
40	33.2	13.2	101912	9	AC000100	AC000100 Homo sapi	
41	33.2	13.2	117937	2	AL139146	AL139146 Homo sapi	
42	33.2	13.2	120951	9	AC073649	AC073649 Homo sapi	
43	33.2	13.2	145383	9	AC010682	AC010682 Homo sapi	
44	33.2	13.2	149794	2	AC007273	AC007273 Homo sapi	
45	33.2	13.2	161656	33	AC027106	AC027106 Homo sapi	

ALIGNMENTS

RESULT: 1
AK024970
LOCUS: AK024970
DEFINITION: Homo sapiens cDNA: FLJ21317 fls, clone COL02275, highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CACCL) mRNA.
ACCESSION: AK024970
VERSION: AK024970.1 GI:10437397
KEYWORDS: oligo capping; fls (full insert sequence).
SOURCE: Homo sapiens
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE: 1 (sites)
Kawabata, A., Hiki, T., Kobateke, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
TITLE: JOURNAL
REFERENCE: 2 (bases 1 to 2022)

AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shishikura, T., Tanaka, T. and Nakamura, Y.

TITLE Direct Submission

JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokawa-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ids.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source 1. 2022

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="COL02275"

/clone_1lb="COL"

/tissue_type="colon"

/note="cloning vector pME18SFL3"

1. 2022

/note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA"

misc_feature

1. 2022

/note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA"

BASE COUNT 612 a 472 c 453 g 485 t

ORIGIN

Query Match 100.0%; Score 252; DB 9; Length 2022;

Best local Similarity 100.0%; Pred. No. 1.3e-60;

Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGAATTGTGTGTTAGTCTTGAACAATCTGGAACATGCGACTGTAACCGCTC 60

DB 123 CAAGAATTGTGTGTTAGTCTTGAACAATCTGGAACATGCGACTGTAACCGCTC 182

QY 61 AATGAGTGAATCAAGAGGCGACCTTTCTGCTGAGAGAGTGAAGTGGGTCTG 120

DB 183 AATGAGTGAATCAAGAGGCGACCTTTCTGCTGAGAGAGTGAAGTGGGTCTG 242

QY 121 GTTGGAGTGTGACATTTGACAGTGTGCTGATGACAACTGATCAATACAGATAAC 180

DB 243 GTTGGAGTGTGACATTTGACAGTGTGCTGATGACAACTGATCAATACAGATAAC 302

QY 181 AGTGGCAGTGCAGGAGGACACTGCGCCAAAGATTACTGAGAGAGTTCAGAGGAGC 240

DB 303 AGTGGCAGTGCAGGAGGACACTGCGCCAAAGATTACTGAGAGAGTTCAGAGGAGC 362

QY 241 TCCATCTGCAGC 252

DB 363 TCCATCTGCAGC 374

RESULT 2

AF127036 2826 bp mRNA PRI 10-AUG-1999

LOCUS Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)

DEFINITION mRNA, complete cds.

AF127036

AF127036.1 GI:4585468

ACCESSION

VERSION

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2826)

Agnel, M., Vernet, T. and Culouscou, J. M.

Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea

JOURNAL FEBS Lett. 455 (3), 295-301 (1999)

99364503

10437792

PUBMED

REFERENCE 2 (bases 1 to 2826)

Agnel, M., and Culouscou, J. M.

TITLE Direct Submission

JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrieres, Rueil-Malmaison 92500, France

FEATURES

source 1. 2826

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="small intestine; colon"

1. 2826

/gene="CaCC1"

5. 2749

/gene="CaCC1"

/note="bovine epithelial chloride channel homolog"

/codon_start=1

/product="calcium-activated chloride channel protein 1"

/protein_id="AAD25487.1"

/db_xref="GI:4585469"

/translation="MGPKRSVFFILIHLEGLANSILQNNNGEIVVAIDPNVP
EDPILQIQKDWYQASLYLPATCKRPFKRVAVILPPTKTRADYPRKLETKYKNA
DVLVAESTPGNDEPTEOMNGEGERIHLPPIFLAKKLAIEGPGRAFVHEMAH
LRMGVFEVYNDKEFTYLSNGRLQAVRCSGILGTNVNKKCGGSCYTRKCTFNKVTGL
YKSGEFLVQSDTERKASIMFAQHVDSIYEFCTEDNNHKEAPNKNQCNLSTWEVI
RSEDEPKTTPMTQPNPTFSLIDIGRIVCLVDKSGSATNGRLNLNAGOLF
LOTVELGSMVAVTQDSMAHVQSEIQLINSGEDRDLARLPAALSGSGSGLRSA
PVIRKRYPTDSEIVLITDGEDNTISGCFNFKOSGAIHIVVALGPSAABOELSELK
MTGCIQTAVSDVQNNGLIDAFGLASSGNGAVSORSIOESGCLTLQNSQNNQVIV
DSVQKDYLFLLTWTQTPQIILMDPSGOKGFVVDKTKAVLIQIPILAVGTWKY
SLQASQILTLTVTSRASNATLPITVTSKTRKDSKPSPLVAVANIRQVAGSPILRA
SVTALIESVNGKTVLELLDNGAGADATDDSDYSSEYFTYDNGRYSKVVALGVA
AARRVYIPQOOSGALYIPGWIENDEIQMNPREDINKDVOHQVCFSSSGSFVAS
DYPNAPIPDLPPPGQITDLKARLHGGSLINLMTAPGDDYDHGTAKYIIRISTSLD
LNDKRESIQVNTVALIPKANSSEYFLEKPNITPENGDTGLFALIQAVDKYDLKSEI
SNARVSLTIPQTPPETSPDETSAFCPNHININSTIIPQIHILKIMKWTGSLQSLIA

BASE COUNT 875 a 623 c 632 g 696 t

ORIGIN

Query Match 100.0%; Score 252; DB 9; Length 2826;

Best local Similarity 100.0%; Pred. No. 1.3e-60;

Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGAATTGTGTGTTAGTCTTGAACAATCTGGAACATGCGACTGTAACCGCTC 60

DB 914 CAAGAATTGTGTGTTAGTCTTGAACAATCTGGAACATGCGACTGTAACCGCTC 973

QY 61 AATGAGTGAATCAAGAGGCGACCTTTCTGCTGAGAGAGTGAAGTGGGTCTG 120

DB 974 AATGAGTGAATCAAGAGGCGACCTTTCTGCTGAGAGAGTGAAGTGGGTCTG 1033

QY 121 GTTGGAGTGTGACATTTGACAGTGTGCTGATGACAACTGATCAATACAGATAAC 180

DB 1034 GTTGGAGTGTGACATTTGACAGTGTGCTGATGACAACTGATCAATACAGATAAC 1093

QY 181 AGTGGCAGTGCAGGAGGACACTGCGCCAAAGATTACTGAGAGAGTTCAGAGGAGC 240

DB 1094 AGTGGCAGTGCAGGAGGACACTGCGCCAAAGATTACTGAGAGAGTTCAGAGGAGC 1153

QY 241 TCCATCTGCAGC 252

DB 1154 TCCATCTGCAGC 1165

RESULT 3

AX193489 3311 bp DNA PAT 15-AUG-2001

LOCUS AX193489

DEFINITION Sequence 1056 from Patent WO0149716.

AX193489

VERSION AX193489.1 GI:15211440
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3311)
AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J., Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and methods for their use
JOURNAL Patent: WO 0149716-A 1056 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN
Query Match 100.0%; Score 252; DB 6; Length 3311;
Best Local Similarity 100.0%; Pred. No. 1.2e-60;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGAATTGTGTGTTAGTCTGTCGCAAAATCTGGAAGCATGCGACTGTAACCGCTC 60
DB 1261 CAAGAATTGTGTGTTAGTCTGTCGCAAAATCTGGAAGCATGCGACTGTAACCGCTC 1320
QY 61 AATCGACTGATCAAGCAGGCGAGCTTTCTGCTGTCGACACAGTGTAGCTGGGCTCTG 120
DB 1321 AATCGACTGATCAAGCAGGCGAGCTTTCTGCTGTCGACACAGTGTAGCTGGGCTCTG 1380
QY 121 GTTGGGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 1381 GTTGGGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 181 AGTGGCAGTGCAGGAGCAGCAGCTCGCCAAAGATTACTGTCAGCAGCTTCAGAGGAGG 240
DB 1441 AGTGGCAGTGCAGGAGCAGCAGCTCGCCAAAGATTACTGTCAGCAGCTTCAGAGGAGG 1500
QY 241 TCCATCTGCAGC 252
DB 1501 TCCATCTGCAGC 1512
RESULT 4
LOCUS AF039400 3311 bp mRNA PRI 14-DEC-1998
DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCICL1) mRNA,
complete cds.
ACCESSION AF039400
VERSION AF039400.1 GI:4009457
KEYWORDS human.
SOURCE
ORGANISM human.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Gruber,A.D., Elble,R.C., Ji,H.L., Schreier,K.D., Fuller,C.M. and Pauli,B.U.
TITLE Genomic cloning, molecular characterization, and functional analysis of human CICL1, the first human member of the family of Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 3311)
AUTHORS Gruber,A.D., Elble,R. and Pauli,B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES
source Location/Qualifiers
1..3311

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/tissue_type="small intestine"
1..3311
/gene="hCICL1"
352..3096
/gene="hCICL1"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC93428.1"
/db_xref="GI:4009458"
/translation="MGPEKSYFILLILHLEGALSNSLIOLNNNGEIVAVADPNVP
EDETLIQIKDVTQASLYLPEATGKREFFKAVALLIPETTKADYRPRKTEYKNA
DVLVAESPGRGDEPYROMCKGKREKFIHTPTPIAGKRLAEVGRCKAVHEKHA
LKMVFDEINDEKFTLSNGRIQAVRCSAGITGIVNVRKCGGSCYTKRCKTNYGL
YKGGCEFLVQSHQTEKASIMFAQHVDSIVETCEONHNKKNOKNOKNLMSTBEVI
RDSDEKFTPTPTQPNPTFSLQIGRIYGLVLDKSGSNATGRLNRLNAGOLF
LOTVELGSWGWTFDSAAHVOSELIOINSGSDRDTLAKRLPAASGGTSISGLSA
FTVIRKKYPTDSESLVLTGDEDMTISCFENKOGSALITHTVALGPSAOLFELELS
MTGGLQYASDVQVONNGUIDARGLSSNGAYSOSIOLESGLTLONSQNNGTIV
DSTVGKDTLELTWTTPQIILMPDSGKGGEVVDKTKAAYLQIPETAKVGTWKY
SLQASSQTLTTLVTSRASNATLPITVTSKTRKDSKPSPLVYANIRQASPLIRA
SVTALIESNGKVTLELDNGAGADATDDQVYSRYFTYDNGRSYKVALGYN
AARRVITPOSCALYIPGWIENDELOMNPREDIKDVOHQVOCFSRTSGSFVAS
DVNNAPIPLPFGSOLTDIKAETIHGGSLLINTPTWAGDQDHGTAKHYIYRSTSLD
LQDKNSLEIQVNTTALIKREANSSEVFLFKPNTITFENCTDLFIQAAYDKVLDSEI
SNIAVRSLETPOTPEPTSPDETAPCPNHNINSTPEIHLTKIMWIGELQSLIA

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN
Query Match 100.0%; Score 252; DB 9; Length 3311;
Best Local Similarity 100.0%; Pred. No. 1.2e-60;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGAATTGTGTGTTAGTCTGTCGCAAAATCTGGAAGCATGCGACTGTAACCGCTC 60
DB 1261 CAAGAATTGTGTGTTAGTCTGTCGCAAAATCTGGAAGCATGCGACTGTAACCGCTC 1320
QY 61 AATCGACTGATCAAGCAGGCGAGCTTTCTGCTGTCGACACAGTGTAGCTGGGCTCTG 120
DB 1321 AATCGACTGATCAAGCAGGCGAGCTTTCTGCTGTCGACACAGTGTAGCTGGGCTCTG 1380
QY 121 GTTGGGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 1381 GTTGGGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 181 AGTGGCAGTGCAGGAGCAGCAGCTCGCCAAAGATTACTGTCAGCAGCTTCAGAGGAGG 240
DB 1441 AGTGGCAGTGCAGGAGCAGCAGCTCGCCAAAGATTACTGTCAGCAGCTTCAGAGGAGG 1500
QY 241 TCCATCTGCAGC 252
DB 1501 TCCATCTGCAGC 1512
RESULT 5
LOCUS AF039400 35278 bp DNA PRI 14-DEC-1998
DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCICL1) gene,
complete cds.
ACCESSION AF039400
VERSION AF039400.1 GI:4009459
KEYWORDS human.
SOURCE
ORGANISM human.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 35278)

AUTHORS Gruber,A.D., Eible,R.C., Ji,H.L., Schreuer,K.D., Fuller,C.M. and Pauli,B.U.
TITLE Genomic cloning, molecular characterization, and functional analysis of human hCLCA1, the first human member of the family of Ca2+-activated Cl- channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 35278)
AUTHORS Gruber,A.D., Eible,R. and Pauli,B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Department of Pathology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES
source location/Qualifiers
1..35278
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
1..1617
/gene="hCLCA1"
1..33522
/gene="hCLCA1"
767..1265
/rpt_family="L1 transposable element"
/rpt_type="dispersed"
1587..1596
/gene="hCLCA1"
/note="putative"
1618..1708
/gene="hCLCA1"
/number=1
join(1618..1708,1962..2383,6667..6807,6979..7126,6966..9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33522)
/gene="hCLCA1"
/product="calcium-dependent chloride channel-1"
1962..2383
/gene="hCLCA1"
/number=2
join(2222..2383,6667..6807,6979..7126,6966..9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33310)
/gene="hCLCA1"
/codon_start=-1
/product="calcium-dependent chloride channel-1"
/db_xref="GI:4009460"
/translation="MGPFKSSVFLILHLEGLSNSLIQLNNGYEGIVVAIDPNVP
EDFTLIDQIDKMTQASLYLFATGKRFYKKNAILIPETWKTADYVRKLEYYKNA
DLVAESTPGNDPEYTEOMNGEGERIHLPDFJAGKKLAFYGGQAFYVEMH
LRNGVPEYNNDEKFTYLSNGRIQAVRCSAGITGNVYKKGQSCYRCPFNKVTGL
YKCECFYLSRQTERASIMPAQVDSIVFCTBQNNKNAKPNQKCNLRWEYI
RDSDEFKRTPTPTPPNPTFSLQIQRTVCLVLDSSMATGNRLRNQAGLEL
LQVVELGSWGVTFDSAAVQSELIQINSGSDRTIAKRLPAASGSGTICSGLRSA
FTVIRKKYPTDGEIVLITDGEVNTISGCEVYKQSGAILITVALGPSAOLEELSL
MTGLOIYASDOVONNGLIDAFALSSGASORSRIHESKGLITLQNSOMNCTVY
DSYTKDTFLITMTQPOIILMDPSGQSGGVNKNRMKMAIOLIGIAKGTWEL
SLQASSQTLITLTVTSRASNTPITPTSTKNTKSPSPPIVYATNRCASPIILIA
SVTALIESVNGKTYTELELDNGAGADTKDGYTSKFTTIDIGRSVYRATGCVN
AARRRVTPOQSGALYIPGMLENDEIQNNPPPEIINKDVQHKQVCFSTSSGSFVAS
DVNPAPIPDPEPGQITDLNAELHGGSLINLTWAPDDVDHGAHRYIIRISYIIL
LRDFNSLQVNTALIPKANSSEVFLFPEVNTFENGCDLFAIADAVKVDKSEI
SNLARVSLFIPQPTPEPSPDERSACPNHINSTIPGIHLIKIMKVIQELQLSIA
"

exon 6667..6807
/gene="hCLCA1"
/number=3
6979..7126
/gene="hCLCA1"
/number=4
9696..9801
exon

BASE COUNT 11245 a 6870 c 7230 g 9933 t
ORIGIN
Query Match 83.08; Score 209.2; DB 9; Length 35278;
Best local similarity 98.6%; Pred. No. 9, 6e-49;
Matches 211; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CY 39 CATGGCAGCTGTAAACCGCCATCATGACATGAATCAAGCAGCCAGCTTTTCCGTGCA 98
DB 19776 CATTAGACGTGTAAACCGCCATCATGACATGAATCAAGCAGCCAGCTTTTCCGTGCA 19835
CY 99 GACAGTTGAGCTGGGCTCTGGTGGATGTCACATTTGACAGTCTGCCATGTACA 158
DB 19846 GACAGTTGAGCTGGGCTCTGGTGGATGTCACATTTGACAGTCTGCCATGTACA 19895
CY 159 AAGTGAACCTATACAGATTAACAGTGGCAGTGCAGAGGACACATCGCCAAAGATTAC 218
DB 19846 AAGTGAACCTATACAGATTAACAGTGGCAGTGCAGAGGACACATCGCCAAAGATTAC 19955
CY 219 TGCAGCAGCTTCAGAGGAGGAGCTCATCTGCAGC 252
DB 19936 TGCAGCAGCTTCAGAGGAGGAGCTCATCTGCAGC 19989
RESULT 6
AL358950 164891 bp DNA HTG 23-JAN-2001
LOCUS Homo sapiens chromosome 1, clone RP11-444C12, *** SEQUENCING IN
DEFINITION PROGRESS ***; 10 unordered pieces.
ACCESSION AL358950
VERSION AL358950.4 GI:12539689
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 164891)
AUTHORS Plumb,B.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,

COMMENT

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 26, 2001 this sequence version replaced gi:9988471.

Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

Project Information

Center project name: b444C12

Summary Statistics

Sequencing program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 161347 bases at least Q40

Consensus quality: 162610 bases at least Q30

Consensus quality: 163328 bases at least Q20

Insert size: 163921; sum-of-contigs

Insert size: 163373; 6.6% error; agarose-fp

Quality coverage: 5.19x in Q20 bases; sum-of-contigs quality coverage: 5.21x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 34236: contig of 34236 bp in length
* 34237 34336: gap of 100 bp
* 34337 47372: contig of 13036 bp in length
* 47373 47472: gap of 100 bp
* 47473 67025: contig of 19553 bp in length
* 67026 67125: gap of 100 bp
* 67126 71239: contig of 4114 bp in length
* 71240 71339: gap of 100 bp
* 71340 83903: contig of 12564 bp in length
* 83904 84003: gap of 100 bp
* 84004 91357: contig of 7354 bp in length
* 91358 91457: gap of 100 bp
* 91458 96452: contig of 4995 bp in length
* 96453 96552: gap of 100 bp
* 96553 148189: contig of 51637 bp in length
* 148190 148289: gap of 100 bp
* 148290 159778: contig of 11489 bp in length
* 159779 159878: gap of 100 bp
* 159879 164891: contig of 5013 bp in length.

FEATURES

SOURCE

1. 164891

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/clone="RP11-444C12"

/clone_lib="RPC1-11.2"

1. 34236

/note="assembly_fragment:00027"

fragment_chain:1

clone_end:T7

vector_side:left"

34337..47372

/note="assembly_fragment:00312"

fragment_chain:1"

47473..67025

/note="assembly_fragment:01923"

fragment_chain:1"

67126..71239

/note="assembly_fragment:02094"

fragment_chain:1"

71340..83903

/note="assembly_fragment:02213"

fragment_chain:1"

misc_feature

84004..91357
/note="assembly_fragment:01292"

misc_feature

91458..96452
/note="assembly_fragment:01391"

misc_feature

96553..148189
/note="assembly_fragment:02045"

misc_feature

148290..159778
/note="assembly_fragment:00427"

misc_feature

159879..164891
/note="assembly_fragment:00773"

misc_feature

fragment_chain:1
clone_end:SP6

misc_feature

vector_side:right"

BASE COUNT 48381 a 32179 c 32357 g 51067 t 907 others

ORIGIN

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

Query Match 83.0%; Score 209.2; DB 2; Length 164891;
Best Local Similarity 98.6%; Pred. No. 7.4e-49;
Matches 211; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

39 CATGGCGATGGTGAACCGCTCATGACTGANTCAAGCAGGCGACCTTTCTGCTGCA 98

DB 62227 CATTAGACTGTAACCGCTCATGACTGANTCAAGCAGGCGACCTTTCTGCTGCA 62168

99 GACAGTTGAGCTGGGCTCTGCTGGTGGATGTGACATTTGACAGTGTGCTGCCATGTACA 158

DB 62157 GACAGTTGAGCTGGGCTCTGCTGGTGGATGTGACATTTGACAGTGTGCTGCCATGTACA 62108

159 AAGTGAACATCATACATTAACAGTGTGACAGGAGCACACCTGCCAAAAGATTACC 218

DB 62107 AAGTGAACATCATACATTAACAGTGTGACAGGAGCACACCTGCCAAAAGATTACC 62048

219 TGCAGCAGCTTCAGGAGGAGCTCATCTGCACG 252

DB 62047 TGCAGCAGCTTCAGGAGGAGCTCATCTGCACG 62014

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

Query Match 83.0%; Score 209.2; DB 2; Length 164891;
Best Local Similarity 98.6%; Pred. No. 7.4e-49;
Matches 211; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

39 CATGGCGATGGTGAACCGCTCATGACTGANTCAAGCAGGCGACCTTTCTGCTGCA 98

DB 62227 CATTAGACTGTAACCGCTCATGACTGANTCAAGCAGGCGACCTTTCTGCTGCA 62168

99 GACAGTTGAGCTGGGCTCTGCTGGTGGATGTGACATTTGACAGTGTGCTGCCATGTACA 158

DB 62157 GACAGTTGAGCTGGGCTCTGCTGGTGGATGTGACATTTGACAGTGTGCTGCCATGTACA 62108

159 AAGTGAACATCATACATTAACAGTGTGACAGGAGCACACCTGCCAAAAGATTACC 218

DB 62107 AAGTGAACATCATACATTAACAGTGTGACAGGAGCACACCTGCCAAAAGATTACC 62048

219 TGCAGCAGCTTCAGGAGGAGCTCATCTGCACG 252

DB 62047 TGCAGCAGCTTCAGGAGGAGCTCATCTGCACG 62014

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

Query Match 83.0%; Score 209.2; DB 2; Length 164891;
Best Local Similarity 98.6%; Pred. No. 7.4e-49;
Matches 211; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

39 CATGGCGATGGTGAACCGCTCATGACTGANTCAAGCAGGCGACCTTTCTGCTGCA 98

DB 62227 CATTAGACTGTAACCGCTCATGACTGANTCAAGCAGGCGACCTTTCTGCTGCA 62168

99 GACAGTTGAGCTGGGCTCTGCTGGTGGATGTGACATTTGACAGTGTGCTGCCATGTACA 158

DB 62157 GACAGTTGAGCTGGGCTCTGCTGGTGGATGTGACATTTGACAGTGTGCTGCCATGTACA 62108

159 AAGTGAACATCATACATTAACAGTGTGACAGGAGCACACCTGCCAAAAGATTACC 218

DB 62107 AAGTGAACATCATACATTAACAGTGTGACAGGAGCACACCTGCCAAAAGATTACC 62048

219 TGCAGCAGCTTCAGGAGGAGCTCATCTGCACG 252

DB 62047 TGCAGCAGCTTCAGGAGGAGCTCATCTGCACG 62014

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

Query Match 83.0%; Score 209.2; DB 2; Length 164891;
Best Local Similarity 98.6%; Pred. No. 7.4e-49;
Matches 211; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

39 CATGGCGATGGTGAACCGCTCATGACTGANTCAAGCAGGCGACCTTTCTGCTGCA 98

DB 62227 CATTAGACTGTAACCGCTCATGACTGANTCAAGCAGGCGACCTTTCTGCTGCA 62168

99 GACAGTTGAGCTGGGCTCT

RESULT 10
AF127035

Dy 11-21 GTTGGGATGGTGACATTTCAGCTGCTGCCCATGTACAAAGTGAACCTCATTACAGATTAAAC 180
Dp 11-21 GTTGGGATGGTGACATTTCAGCTGCTGCCCATGTACAAAGTGAACCTCATTACAGATTAAAC 180

OY	181	AGTCGACGCAGCAGCACACTGCCCAAAAGAATTACTGGACGAGCTTCCAGAGCGAGC	240
OY	181		
DB	1118	AGCACTGATGAAGAAACACACTCATGCAGAGATTACTATATCTCTGTGGAGAACT	1177
OY	241	TCCATCTGC	249
OY	241		
RESULT	11		
AK000072			
LOCUS	AK000072	3221 bp	mRNA
DEFINITION	Homo sapiens CDNA FLJ20065 fts, clone COL01613, highly similar to ECJC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.		PRI 22-FEB-2000
ACCESSION	AK000072		
VERSION	AK000072.1	GI:7019922	
KEYWORDS	oligo capping; fts (full insert sequence).		
SOURCE	Homo sapiens colon CDNA to mRNA, clone_11b.COL clone:COL01613.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (sites)		
REFERENCES	Kawabata,A., Hikiji,T., Kobateke,N., Inagaki,H., Ikema,Y., Okamoto,S., Ohtani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.		
JOURNAL	NEDO human cDNA sequencing project		
REFERENCE	Unpublished (2000)		
AUTHORS	2 (bases 1 to 3221)		
TITLE	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.		
JOURNAL	Direct Submission		
COMMENT	Submitted (15-FEB-2000) to the DDBJ/EMBL/genBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Vitrology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)		
FEATURES	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Deparment of Vitrology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).		
source	Location/Qualifiers		
	1..3221		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="COL01613"		
	/clone_11b="COL"		
	/tissue_type="colon"		
	/note="cloning vector pMERISFL3"		
	1..3221		
	/note="highly similar to ECJC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"		
BASE COUNT	1105 a 600 c 634 g 882 t		
ORIGIN			
Query Match	55.6%; Score 140.2; DB 9; Length 3221;		
Best Local Similarity	72.7%; Pred. No. 3.8e-29;		
Matches 181; Conservative	0; Mismatches 68; Indels 0; Gaps 0;		
OY	1	CAAAATAATTGTGTGTTTGCTTCACAAATGGAAAGCATGGGACGTGTGTAACGCCCTC	60
DB	952	CAAAATAATTGTGTGTTTGCTTCATATAGTCTGGAGACATGGGGGTTAGACACGCCCTA	1011
OY	61	AATGCAATGAATCAACAGCAGCCAGCTTTTCCTGCTGACAGACAGTTGAGCTGGGGCTCG	120
DB	1012	AATGCAATGAATCAACAGCAGCAAAACATTTCTCGTGCAGACAGCTTGAATAATGGATCTGG	1071
OY	121	GTTGGAGTGACATTGACAGTGCTGCCCATGTGACAAGTGAACATCATACAGATTAAAC	180

```

D5 1072 GTGGGGAGTGGCTTCACTTGGATAGTAAGTACGCCACATTTGTAATAAGCTAATCCAAATTAATA 1131
QY 131 AETGGCAGTGCAGGAGACACACTCGCCAAAGATTACCTGCGACGACCTTCAGAGGGAGC 240
D5 1132 AGCACTGATGAAAGAAACACACTCATGTGCGAGGATTACCTACATATCTCTGGGAGGAACT 1191
QY 241 TCCATCTGC 249
D5 1192 TCCATCTGC 1200
RESULT 12
AX092338 3265 bp DNA PAT 21-MAR-2001
LOCUS AX092338
DEFINITION Sequence 69 from Patent WO0116318.
ACCESSION AX092338
VERSION AX092338.1 GI:13444483
KEYWORDS
SOURCE human;
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3265)
Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
Wood,W.I.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
Patent: WO 0116318-A 69 08-MAR-2001;
JOURNAL Genetech, Inc. (US)
FEATURES
Source Location/Qualifiers
1..3265
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1159 a 596 c 632 g 878 t
ORIGIN
Query Match 55.6%; Score 140.2; DB 6; Length 3265;
Best Local Similarity 72.7%; Pred. No. 3.8e-29;
Matches 181; Conservative 0; Mismatches 68; Indels 0; Gaps 0
QY 1 CAAGAATTTGTGTGTTTACTTCCTGCAAAATCTGGAAGCATGGCGGCTGTAAACGCCCTC 60
D5 534 CAAGATTTGTGTCTTACTTCTTGTAAGTCTGGAAGCATGGGGGTAAAGACCGCTTA 993
QY 61 AATGCAGTAATCAAGCAGCAGCAGCTTTTCTGCTGCAGACAGTTAGCTGGGGTCTCG 120
D5 594 AATGCAGTAATCAAGCAGCAGCAGCAAAACATTTCCGCTGCACAGCTGTGAATAATGATCTCG 1053
QY 121 GTTGGAGTGTACATTGTACAGTGTCTGCCATGTACAAAGTAGAACTATACAGATAAC 180
D5 1054 GTGGGAGTGTTCACATTGTGATAGTACTGCCACATATGTAAATTAAGCTAATCCAAATTAATA 1113
QY 131 AATGGCAGTGCAGGAGCAGCAGCTCGCCAAAGATTACCGCGACGACCTTCAGAGGGAGC 240
D5 1114 AACAGGATGAAGAAAGAAACACACTCATGCGAGGATTACCTACATATCTCTGGGAGGAACT 1173
QY 241 TCCATCTGC 249
D5 1174 TCCATCTGC 1182
RESULT 13
AF047838 3022 bp mRNA ROD 26-NOV-1998
DEFINITION Mus musculus calcium-sensitive chloride conductance protein-1
(musclal) mRNA, complete cds.
ACCESSION AF047838
VERSION AF047838.1 GI:3925280
KEYWORDS house mouse.
SOURCE Mus musculus

```


REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi;
AUTHORS	1 (bases 1 to 3022)	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE	Gandhi,R., Elble,R.C., Gruber,A.D., Schreuer,K.D., Ji,H.L., Fuller,C.M. and Pauli,B.U. Molecular and functional characterization of a calcium-sensitive chloride channel from mouse lung	
JOURNAL	J. Biol. Chem. 273 (48), 32096-32101 (1998)	
REFERENCE	2 (bases 1 to 3022)	
AUTHORS	Gandhi,R., Elble,R.C., Gruber,A.D. and Pauli,B.U.	
JOURNAL	Direct Submission Submitted (12-FEB-1998) Vet. Pathology, Cornell Univ., WMC C4-146, Ithaca, NY 14856, USA	
SOURCE	location/qualifiers 1..3022 /organism="Mus musculus" /db_xref="taxon:10090" /tissue_type="whole lung" 1..3022 /gene="mclca1" 18..2726 /gene="mclca1"	
gene	/note="calcium-activated chloride channel; Alternative chloride channel to CFTR that may influence severity of cystic fibrosis in different tissues; mCAC-1" /codon_start=1 /product="calcium-sensitive chloride conductance protein-1." /protein_id="AAC79982.1" /db_xref="GI:3925281" /translation="MPVGLQVLFLTLHLONTSSWHLNSNGEYGVAINSPVEPDERLIPSIKEKVTASTYLEFASOGARYFRNLSILVPMTKSKPEYLKRSDKALAVIADPHLOHGDDPYLTLOYGGCGRGROYIHPTPNLLIDNRITGPRGRVHMAHI RWCVDDEYNDPOPPMYRKNRIETARSTRITGNVVYHCRCNCVRACRDSDRTIR YEKCTFTPIDKTQTAGASIMPMQNLYSEECTCKNHAEAPNLONKCNRRSTWDYTKTSADFENAMPKGTGEAPPEPTFSLSRRRVCLVLDRSGSMKEDRLIMNOAAELV LINDIVERESNWGVLTPEDSAHIONTYLKITSRSSDYOKITANLPQOASGSICHGLD AGPAITSSDSOSTSGSEIVILLTDDEDNIGIRSCFEVASCSAIIHTIALGPSARELETFY LSDATGGLRFYANKDLNSLIDAFRSISTSSVSQAOLALSFADVAAGAIANGTVAF HPLDJOEDEFVNATVIWMVKRPRIIIOODPKKKYTISPDODKLNTSRSLVIOIPGAETG TWSYSTGTSQSOLTMMVTYTPPARSPMEPIPLATAHMSGSTOYSRMIVAYROSGEPF PVLCANTALILEAHGHQVTELEMDNGADTYANDGITTTFFDYHENGRTSLKVRV OAOGNKTRLSLRQNKSLYIPGYENGLVINPDPVOEAILEATVEDERIVTSGGS FTVSGAPPDGHARVFPSPKVTLDLAERTIGDYIHLTWAPKVLVDNGRAHYIIRMSCG HPILDJOEDEFVNATLVNASLIPKEAGSGETFKPEFEFKIANGLIQLYIAIODNEAST TSESNIAOA VKLTLSBDSISALSALDISAISMTIMGTLVIPNSILN"	
CDS		
BASE COUNT	923 a 710 c 661 g 728 t	
ORIGIN		
Query Match	43.3%; Score 109; DB 10; Length 3022;	
Best Local Similarity	65.3%; Pred. No. 2.4e-20;	
Matches 160; Conservative	0; Mismatches 85; Indels 0; Gaps 0;	
OY	5 GAATGTGCTGTTACTCCTTGACAATAATGGAGAAGCATGGCACTGTAAACCCCTCAATC	64
Dd		996
OY	65 GACTGCATCAAGCAGCGCAGCTTTTCCTGCTGCACACAGTAGTTAGCTGGGGTCTGGGTTG	124
Dd		1056
OY	997 GAATGATCAACGACGACAGCACTGTACTTAACCAATATGTGAAAAAGAGCTATTAGGTTG	1056
OY	125 GGATGTGTACATTTTACAGAGTGTGCCCATGTTCACAAAGTGAAGTCACTATACAGATAAACAGTG	184
Dd		1116
OY	1057 GATTAGTACACATTTACAGCGCTGCCACATCCAAAATTTATCTAATTAATAAAATACGAGTA	1116
OY	185 GCATGACAGGAGACACATCGCCCAAATAATTCCTGCAGCAGCAGCTTCAGAGGAGAGCTCCA	244
Dd		1176
OY	1117 GTAGTGACTACCACAAAGATCACCGCAACCTCCCCAACAGGCTTCTGTGGAACTTCAA	1176
OY	245 TCCTGC 249	

```

Db      1177 TTGC 1181

RESULT 14
LOCUS   AF052746
DEFINITION Mus musculus chloride channel cacc mRNA, complete cds.
ACCESSION AF052746
KEYWORDS AF052746.1 GI:3560546
SOURCE  house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3471)
AUTHORS  Romio,L., Musante,L., Cintl,R., Serl,M., Moran,O., Zagarra-Moran,O.,
and Galietta,L.J.V.
TITLE   Characterization of a murine gene homologous to the bovine Cacc
chloride channel
JOURNAL  Gene 228 (1-2), 181-188 (1999)
MEDLINE  99173882
REFERENCE 2 (bases 1 to 3471)
AUTHORS  Romio,L., Musante,L., Cintl,R., Moran,O., Serl,M. and
Galietta,L.J.V.
TITLE   Direct Submission
JOURNAL  Submitted (09-MAR-1998) Lab. di Genetica Molecolare, Istituto
Glennina Gaslini, Genova I-16148, Italy
FEATURES
SOURCE
1..3471
/organism="Mus musculus"
/db_xref="taxon:10090"
/db_xref="dbEST:AA245677"
47..2752
/codon_start=1
/product="chloride channel cacc"
/protein_id="AAC35003.1"
/db_xref="GI:3560547"
/translation="MVPGLQYLFLTLHLQNTESMWHLNSGEGVAINPSPVE
DEIRLPSIKEMWQASTYLFEASQGRVFNILSLVPMWMSKREYLPMPKESYDKAD
VIVLADPHLOHDDDDYTLQYGGCGRGQYIHPTPNLILNDLRIYGPGRVHVMHML
RMGVEEDYNNDDPPEYMSRKNTLETRGTSRTGRTGNNHNGRCNGVIRACGRDSTKL
YERKCTPIPKDTIQTAGASINPMQNLNSVECTEKNHNAEPNLQNKRRSTWDYI
KTSDFQNPAPRGTGEAPPEPTSLLSRRRVVCLVDKSSMKERLIMNOAAEL
VLYDIVERKESWVGIVTFDSAHIQNTYILKITSSDYOKITANLPQASGGSTCHGLQ
AGPGLATISSDQSTGSSEIVLTLDEGDNGIRSCFEAVSRSGAIHPTIALGPSARELETP
LSDMTGGLRFYFANKDLIDAFSRISSTSGVSQOALLESKAFDVRAGMINGTVB
LSDVNGDITFEVITMWVKKEPIILDDPKRKYTSDPODKLIRSAVRIQIPGAEFG
TWYTSITGTSQLTMTVTTPRASPMPLEPILALAHMSQSPQVSRMIVAYRVSQGL
PVLCAANTALILEAHGHQVTELEMDNCGADTVANDGITYTFDTHGNSRYSLKRV
QAOQNKTRLRLQNKNSLXIPGYENKIVLNPRPDVQEAIDATVEDRFRTVSGSS
FTYSGAPPDGDAHVAFPPSKVTLDEAFETIGYIHLITAPKVLVDNGRAHRYILRMSQ
HPDLQEDFNNATLVNASSILPKFEGAKRETFKPEFEFKIANGIOLYIAQADNEASEL
TSEVSNIAQAVKTLSEDSISALDIDGIAICMTIWLIVIFNSILN"
BASE COUNT 1096 a 777 c 736 g 862 t
ORIGIN
Query Match 43.3%; Score 109; DB 10; Length 3471;
Best local Similarity 65.3%; Pred. No. 2.3e-20;
Matches 160; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```


Db 476 ATGTTTATATCAAGTGAACAAAGTGTGATATATTCCTGTGTTTCGATTTCAGACAG 535
QY 240 GAAATGAG 248
Db 536 AAGGTTTAC 544

RESULT 2
US-08-083-590A-21/c

Sequence 21, Application us/08083590A
Patent No. 5786158

GENERAL INFORMATION:

APPLICANT: Artaanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York

COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083.590A
FILING DATE: 25-JUN-1993

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:
LENGTH: 9723 base pairs

TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: unknown
MOLECULE TYPE: CDNA

FEATURE:
NAME/KEY: CDS

LOCATION: 10..7419
US-08-083-590A-21

Query Match 9.9%; Score 28.6; DB 1; Length 9723;
Best Local Similarity 49.0%; Pred. No. 6.1;
Matches 76; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 94 TGCATCATCCACACAGTCGCTTTGGGGCCCTTCGACGTCAGAACTGAGAGGCTGTC 153
Db 2221 TGAGACCTCCAGTACAGTTTCATGATGAGGATTTGCTCAGGCAATTCCTTACCTGTG 2162

QY 154 CAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTTCAGAACAAATGGCTTCAT 213
Db 2161 AATAGCAGCTGGGGTGTATGGGGTCCCTCGGGGCAATATACGCGAAACCATTCACACCGT 2102

QY 214 TGATGCTTTGGGGCCCTTCATCAGGAATGAG 248
Db 2101 TGATACATGTTGCAACCTTCGACAGGAGGATTTGAG 2067

RESULT 3
US-08-532-384-21/c

Sequence 21, Application us/08532384
Patent No. 6083904

GENERAL INFORMATION:

APPLICANT: Artaanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York

COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,384

FILING DATE:
CLASSIFICATION: 424

PREVIOUS APPLICATION DATA:
APPLICATION NUMBER: 08/083,590

FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:

NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:
LENGTH: 9723 base pairs

TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: unknown
MOLECULE TYPE: CDNA

FEATURE:
NAME/KEY: CDS

LOCATION: 10..7419
US-08-532-384-21

Query Match 9.9%; Score 28.6; DB 3; Length 9723;
Best Local Similarity 49.0%; Pred. No. 6.1;
Matches 76; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 94 TGCATCATCCACACAGTCGCTTTGGGGCCCTTCGACGTCAGAACTGAGAGGCTGTC 153
Db 2221 TGAGACCTCCAGTACAGTTTCATGATGAGGATTTGCTCAGGCAATTCCTTACCTGTG 2162

QY 154 CAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTTCAGAACAAATGGCTTCAT 213
Db 2161 AATAGCAGCTGGGGTGTATGGGGTCCCTCGGGGCAATATACGCGAAACCATTCACACCGT 2102

QY 214 TGATGCTTTGGGGCCCTTCATCAGGAATGAG 248
Db 2101 TGATACATGTTGCAACCTTCGACAGGAGGATTTGAG 2067

RESULT 4
US-08-594-031-80
Sequence 80, Application us/08594031
Patent No. 5783192
GENERAL INFORMATION:
APPLICANT: THOMPSON, Timothy C.
TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
NUMBER OF SEQUENCES: 175


```

CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 1299 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004-2400

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,031
FILING DATE: 30-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/006,838
FILING DATE: 16-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Remenick, James
REGISTRATION NUMBER: 36,902
REFERENCE/DOCKET NUMBER: 0A146-0110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX:

INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 371 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-594-031-80

Query Match          9.7%: Score 28; DB 1; Length 371;
Best Local Similarity 55.0%: Pred. No. 1.9;
Matches 55; Conservative 0; Mismatches 45; Indels 0; Gaps 0.

QY 155 AAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTACAGAAATGAGCCCTCATT 214
      |||||  |||||||  |  ||  |||  |||||||  |  ||  ||
Db 75 AAAATGCCAATATGTTTACACAGCTGTGAACAGGSGTCAGTTCAAGATCCTGATGGCTTG 134
      |||||  |||||  |  |||||||  |||  |||  |

QY 215 GATGCTTTGGGGCCCTTTCATCAGAGAAATGAGAGCTGCT 254
      ||  ||  ||  |||||||  |||  |||  |
Db 135 AGAATGCTTCTTGTTCACATCAGAGAACTGGAATGTTT 174

RESULT 5
US-08-961-083-127/c
; Sequence 127, Application US/08961083
; Patent No. 6159469
;
GENERAL INFORMATION:
APPLICANT: Chol et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2

```

```

SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/961,083
  FILING DATE:
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:
  FILING DATE:
  ATTORNEY/AGENT INFORMATION:
    NAME: Brookes, A. Anders
    REGISTRATION NUMBER: 36,373
  REFERENCE/DOCKET NUMBER: PB340P2
  TELEPHONE: (301) 309-8504
  TELEFAX: (301) 309-8512
  INFORMATION FOR SEQ ID NO: 127:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 352 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
US-08-961-083-127

Query: Match          9.4%; Score 27.2; DB 3; Length 352;
Best Local Similarity 52.2%; Pted. No. 3.5;
Matches 59; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Dy 1 33 TCGTCAGCAGTGGGAGACACATCATAGTGGTCTTTAACGAGGTCAACCAAGTN 92
   1 111 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Dh 1 319 TACGCTGTGTGATGACTGCCAAATATAGTATCTGCTTCAAGAGGATCTGTGCTCA 280
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OY 1 93 GTGCATCATCCACACAGTCGCTTTGGGCGCCCTCTGCAGCTCAAGACTAGAG 145
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 1 279 AATTCACATCCCAATGGGGAATTTCTCTGCTCGAACCACCAATATTCAG 227
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT: 6
US-08-743-637B-30
Sequence 30, Application US/08743637B
Patent No. 5994066

GENERAL INFORMATION:
APPLICANT: BERGERON, Michel G.
APPLICANT: PICARD, Francois J.
APPLICANT: OUELLETTE, Marc
APPLICANT: ROY, Paul H.
TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHARLES & BRADY
STREET: 411 EAST WISCONSIN AVENUE
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202-4497

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/743,637B
  FILING DATE: 04-NOV-1996
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/526,840
    FILING DATE: 11-SEP-1995
  ATTORNEY/AGENT INFORMATION:
    NAME: BAKER, Jean C.
    REGISTRATION NUMBER: 35,433

```


TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 8133 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..8130
US-08-480-604A-5

Query Match 9.3%; Score 26.8; DB 1; Length 8133;
Best Local Similarity 57.0%; Pred. No. 24;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 173 CAGACATATGCTTCAGATCAAGTTCAGAACATGGCCCTATTGAGCTTTGGGCCCTT 232
DB 1567 CAAGCAAGTGCAGAAATATGATTTGAGAAATATGTAAAGAGATTATCTGGTGCATCTCTT 1626
QY 233 TCATCAGAAATGAGCTGTCTCTCA 258
DB 1627 TCTGAGACATAGGGGTAGACTTTAA 1652

RESULT 9

US-08-405-496A-5
Sequence 5, Application US/08405496A
Patent No. 5919665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 8133 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..8130
US-08-495-496A-5

Query Match 9.3%; Score 26.8; DB 2; Length 8133;
Best Local Similarity 57.0%; Pred. No. 24;
Matches 49; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 173 CAGACATATGCTTCAGATCAAGTTCAGAACATGGCCCTATTGAGCTTTGGGCCCTT 232
DB 1567 CAAGCAAGTGCAGAAATATGATTTGAGAAATATGTAAAGAGATTATCTGGTGCATCTCTT 1626
QY 233 TCATCAGAAATGAGCTGTCTCTCA 258
DB 1627 TCTGAGACATAGGGGTAGACTTTAA 1652

RESULT 10

US-08-915-136-5
Sequence 5, Application US/08915136
Patent No. 6290960

GENERAL INFORMATION:

APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,136
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,604
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.


```

1      REGISTRATION NUMBER: 40, 027
2      REFERENCE/DOCKET NUMBER: OPMD-01763
3      TELECOMMUNICATION INFORMATION:
4      TELEPHONE: (415) 705-8410
5      TELEFAX: (415) 397-8338
6      INFORMATION FOR SEO ID NO: 5:
7      SEQUENCE CHARACTERISTICS:
8      LENGTH: 8133 base pairs
9      TYPE: nucleic acid
10     STRANDEDNESS: single
11     TOPOLOGY: linear
12     MOLECULE TYPE: DNA (genomic)
13     FEATURE:
14     NAME/KEY: CDS
15     LOCATION: 1..8130
16     US-08-915-136-5

```

Query Match	9.3%	Score 26.8	DB 4	Length 8133
Best Local Similarity	57.0%	Pred. No. 24		
Matches 49	Conservative 0	Mismatches 37	Indels 0	Gaps 0

Qy 173 CAGCATATGCTTCAGATCAAGTTGAGAACAATGGCCCTCATTGATGCTTTTGGGGCCCTT 232
|| || ||| | |||| | ||| ||| ||| ||| |||
Db 1567 CAAGCAAGTGCAAAATAATCAATTTGAGAAATATGTAAAGATTATACGTGCGATCTCTT 1626

QY	233	TCATCAGGAATGGAGCTGTCTCTCA	258
Db	1627	TCTGAGACCAATGGGGTAGACTTTAA	1652

```

11 RESULT
12 US-09-338-907-183/c
13 Sequence 183, Application US/09338907
14 Patent No. 6255516
15 GENERAL INFORMATION:
16 APPLICANT: Cohen, Daniel
17 APPLICANT: Blumenfeld, Marta
18 APPLICANT: Ilya, Chumakov
19 APPLICANT: Bougueleret, Lydie
20 TITLE OF INVENTION: PROSTATE CANCER GENE
21 FILE REFERENCE: GENSET 18CP1CP
22 CURRENT APPLICATION NUMBER: US/09/338,907
23 CURRENT FILING DATE: 1999-06-23
24 EARLIER APPLICATION NUMBER: 08/996,306
25 EARLIER FILING DATE: 1997-12-22
26 EARLIER APPLICATION NUMBER: 60/099,658
27 EARLIER FILING DATE: 1998-09-09
28 EARLIER APPLICATION NUMBER: 09/218,207
29 EARLIER FILING DATE: 1998-12-22
30 NUMBER OF SEQ ID NOS: 578
31 SOFTWARE: Patent.pm
32 SEQ ID NO 183
33 LENGTH: 37950
34 TYPE: DNA
35 ORGANISM: Mus musculus
36 FEATURE:
37 NAME/KEY: exon
38 LOCATION: 5259..5328
39 OTHER INFORMATION: exon2
40 FEATURE:
41 NAME/KEY: exon
42 LOCATION: 12675..12791
43 OTHER INFORMATION: exon3
44 FEATURE:
45 NAME/KEY: exon
46 LOCATION: 14621..14710
47 OTHER INFORMATION: exon4
48 FEATURE:
49 NAME/KEY: exon
50 LOCATION: 18922..19912
51 OTHER INFORMATION: exon5
52 FEATURE:

```

```

1 NAME/KEY: exon
2 LOCATION: 21189..21950
3 OTHER INFORMATION: exon6
4
5 FEATURE:
6 NAME/KEY: exon
7 LOCATION: 23877..23510
8 OTHER INFORMATION: exon7
9
10 FEATURE:
11 NAME/KEY: exon
12 LOCATION: 25520..26016
13 OTHER INFORMATION: exon8
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
```

Query Match	9.3%	Score	26.8	DB	4	Length	37950
Best focal Similarity	63.5%	Pred. No.	52				
Matches	40	Conservative	0	Mismatches	23	Indels	0
						Gaps	0

Cy 76 CGAGGTCAACAAGTNGTCATCATCCACACAGTCGCTTTGGGCCCTCGCAGCTCA 135
 | | | | | | | | | | | | | | | |
Nb 7699 CAAAGGCACACAGCCATGCTTCAGCCGCTGTACTGTGAGACCACAGGAGACA 7640

QY	136	AGA	138
D0	7639	AGA	7637

RESULT 12

; Sequence 24, Application US/08659188

1 APPLICANT: Yanofsky, Martin F.
 2 TITLE OF INVENTION: Seed Plants Exhibiting Inducible Early
 3 TITLE OF INVENTION: Reproductive Development and Methods of Making Same
 4 NUMBER OF SEQUENCES: 26
 5 CORRESPONDENCE ADDRESS:
 6 ADDRESSEE: Campbell and Flores
 7 STREET: 4370 La Jolla Village Drive, Suite 700
 8 CITY: San Diego
 9 STATE: California
 10 COUNTRY: USA
 11 ZIP: 92122
 12 COMPUTER READABLE FORM:
 13 MEDIUM TYPE: Floppy disk
 14 COMPUTER: IBM PC compatible
 15 OPERATING SYSTEM: PC-DOS/MS-DOS
 16 SOFTWARE: Patentln Release #1.0, Version #1.25
 17 CURRENT APPLICATION DATA:
 18 APPLICATION NUMBER: US/08/659,188
 19 FILING DATE: 05-JUN-1996
 20 CLASSIFICATION: 800
 21 ATTORNEY/AGENT INFORMATION:
 22 NAME: Campbell, Cathryn A.
 23 REGISTRATION NUMBER: 31,815
 24 REFERENCE/DOCKET NUMBER: P-UD 1946
 25 TELECOMMUNICATION INFORMATION:
 26 TELEPHONE: (619) 535-9901
 27 TELEFAX: (619) 535-8949
 28 INFORMATION FOR SEQ ID NO: 24:
 29 SEQUENCE CHARACTERISTICS:
 30 LENGTH: 371 base pairs
 31 TYPE: nucleic acid
 32 STRANDEDNESS: double
 33 TOPOLOGY: linear
 34 FEATURE:
 35 NAME/KEY: misc_feature
 36 LOCATION: 1..371
 37 OTHER INFORMATION: /note="element = heat shock
 38 OTHER INFORMATION: inducible regulatory element (HSPol-1 promoter)."
 39 US-08-659-188-24

Query Match 9.2%; Score 26.6; DB 3; Length 371;
Best Local Similarity 55.6%; Pred. No. 5.8;
Matches 50; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

DB 14 GATGATCTGAATTTGCTGCTGACGATGGGAGACAACTTAAGTGGTCTTT 73
222 GAGAGATTGTGATTTGTTTGTGATTTATGAAGAGAGAGGAAAAACAAGTATTT 163

QY 74 AACGAGTCAAAACAAGTNGTCCATC 103
162 ATAGAGAGAAATAAGTCTAGTCTCTC 133

DB 162 ATAGAGAGAAATAAGTCTAGTCTCTC 133

RESULT 13
US-08-655-227-24/C
; Sequence 24, Application US/08655227
; Patent No. 6025483

GENERAL INFORMATION:
APPLICANT: Yanofsky, Martin F.
TITLE OF INVENTION: Maize and Cauliflower APETAL1 Gene
TITLE OF INVENTION: Products and Nucleic Acid Molecules Encoding Same
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,227
FILING DATE: 05-JUN-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 2143
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 371 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..371
OTHER INFORMATION: /note="element = heat shock
; OTHER INFORMATION: Inducible regulatory element (HSP81-1 promoter)."
US-08-655-227-24

Query Match 9.2%; Score 26.6; DB 3; Length 371;
Best Local Similarity 55.6%; Pred. No. 5.8;
Matches 50; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

DB 14 GATGATCTGAATTTGCTGCTGACGATGGGAGACAACTTAAGTGGTCTTT 73
222 GAGAGATTGTGATTTGTTTGTGATTTATGAAGAGAGAGGAAAAACAAGTATTT 163

QY 74 AACGAGTCAAAACAAGTNGTCCATC 103
162 ATAGAGAGAAATAAGTCTAGTCTCTC 133

DB 162 ATAGAGAGAAATAAGTCTAGTCTCTC 133

RESULT 14

US-08-655-241-24/C
; Sequence 24, Application US/08655241
; Patent No. 6025543

GENERAL INFORMATION:
APPLICANT: Yanofsky, Martin F.
APPLICANT: Reigel, Delfel
TITLE OF INVENTION: Seed Plants Exhibiting Early Reproductive
TITLE OF INVENTION: Development and Methods of Making Same
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,241
FILING DATE: 05-JUN-1996
CLASSIFICATION: CLASS 800
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 1894
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 371 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..371
OTHER INFORMATION: /note="element = heat shock
; OTHER INFORMATION: Inducible regulatory element (HSP81-1 promoter)."
US-08-655-241-24

Query Match 9.2%; Score 26.6; DB 3; Length 371;
Best Local Similarity 55.6%; Pred. No. 5.8;
Matches 50; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

DB 14 GATGATCTGAATTTGCTGCTGACGATGGGAGACAACTTAAGTGGTCTTT 73
222 GAGAGATTGTGATTTGTTTGTGATTTATGAAGAGAGAGGAAAAACAAGTATTT 163

QY 74 AACGAGTCAAAACAAGTNGTCCATC 103
162 ATAGAGAGAAATAAGTCTAGTCTCTC 133

DB 162 ATAGAGAGAAATAAGTCTAGTCTCTC 133

RESULT 15
US-08-991-789A-35
; Sequence 35, Application US/08991789A
; Patent No. 6225054

GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
APPLICANT: Reed, Steven G.
APPLICANT: Smith, John M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 292
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed IP Law Group
STREET: 701 Fifth Avenue, Suite 6300


```

: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104-7092
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/991.789A
: FILING DATE: 11-Dec-1997
: CLASSIFICATION: <Unknown>
: ATTORNEY/AGENT INFORMATION:
:   NAME: Potter, Jane E. R.
:   REGISTRATION NUMBER: 33,332
:   REFERENCE/DOCKET NUMBER: 210121.419C3
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: (206) 622-4900
:   TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 35:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 578 base pairs
:     TYPE: nucleic acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
:   SEQUENCE DESCRIPTION: SEQ ID NO: 35:
:
: US-08-991-789A-35

```

```

Query Match          9.28; Score 26.6; DB 4; Length 578;
Best Local Similarity 63.18; Pred. No. 7.2;
Matches 41; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 158 ATGACAGAGGTTTACACATATGCTTCATGATCAAGTTTCAAGACATGGGCTCATTTGAT 217
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 379 ACGCCTGGTGTAAAGACTATCTTTGGCCTCAGTTTCAAGATGCGCAAGTGT 438

QY 218 GCTTT 222
   | |
DB 439 TCCTT 443

```

Search completed: April 3, 2002, 20:53:44
 Job time: 41782 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:23 ; Search time 16681 Seconds
(without alignments)
186.171 Million cell updates/sec

Title: US-09-049-696-7

Perfect score: 289
Sequence: 1 GAATATCCACTGATGAT.....ACCTGAGAGTAAGGATTA 289

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estol:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_esti:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vtl:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	268	92.7	657	10	AW361521 OV2-CT026
2	264.6	91.6	625	10	AW361532 OV2-CT026
3	263.2	91.1	621	10	AW361520 OV2-CT026
4	260.2	90.0	654	10	AW361522 OV2-CT026
5	253.2	87.6	652	10	AW753449 OV2-CT026
6	230	79.6	561	10	AW361523 OV2-CT026
7	200.6	69.4	661	11	BF581041 602100173
8	200.6	69.4	2933	12	AK008659 Mus muscu
9	196	67.8	747	11	BF580957 602100659
10	188.2	65.1	704	11	BG965023 602831668
11	188	65.1	2915	12	AK007466 Mus muscu
12	161.2	55.8	439	13	AQ558859 HS_2089_B

13	159.8	55.3	831	11	BG286204
14	159.6	55.2	504	13	AQ557463
15	119	41.2	275	11	BG953500
16	113.2	39.2	744	13	AZ708112
17	109.4	37.9	1033	11	BG243305
18	106.2	36.7	462	10	AI851265
19	101	34.9	510	10	AW762590
20	94.4	32.7	910	11	BI112685
21	94.2	32.6	858	11	BG675746
22	87.4	30.2	550	11	BF355533
23	85.2	29.5	517	12	AK006921
24	84.4	29.2	534	11	BF355513
25	80.8	28.0	940	11	BG678187
26	77.4	26.8	819	11	BG192681
27	77	26.6	468	10	AA726662
28	75.8	26.2	820	11	BG186446
29	70.8	24.5	337	10	AV047614
30	69.4	24.0	435	10	AA443258
31	68	23.5	453	13	AQ207683
32	67.2	23.3	906	11	BF581342
33	67	23.2	437	10	BE382180
34	61.2	21.2	461	13	AZ771825
35	59.6	20.6	455	10	AA871197
36	58.2	20.1	584	11	BF116637
37	55.6	19.2	207	10	AF071903
38	55	19.0	184	13	AZ706870
39	53.8	18.6	866	11	BI155803
40	49.4	17.1	314	10	BB225759
41	47.4	16.4	823	11	BG741082
42	46.2	16.0	331	10	BA499841
43	45.4	15.7	392	13	AQ082322
44	44	15.2	300	10	BB560274
45	41.2	14.3	299	10	BB124752

ALIGNMENTS

RESULT 1
LOCUS AW361521 657 bp mRNA EST 04-FEB-2000
DEFINITION OV2-CT0261-261099-011-f03 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW361521
VERSION AW361521.1 GI:6866275
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 657)
HCGP <http://www.ludwig.org.br/ORESTES>.
TITLES The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
CONTACT: Simpson A.J.G.
COMMENT: Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&ct=QV2-CT0261-261099-011-f03&ct3=1999-10-26&ct4=1>)
Seq primer: puc 18 forward
High quality sequence start: 123
Location/Qualifiers
1. 657
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="CT0261"

FEATURES


```

/dev storage="Adult"
/note="Organ: colon; Vector: puc1b; Site_1: Sma1; Site_2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      175 a      160 c      164 g      158 t
ORIGIN

```

	Query Match	Similarity	92.7%	Score	268;	DB	10;	Length	657;
	Best Local	Similarity	97.6%	Pred.	No.	1,	le=68;		
	Matches	282;	Conservative	0;	Mismatches	6;	Indels	1;	Gaps
QY	2	AAATATCCAACTGATGGA-TCTGAAATTTGCTGCTGACGATGAGGGAAGACACACTAT	60						
Db	90	AAATTTCTTAACGATGGATTCGTAAATGTTGCTGCTGACGAGATGGGAAGAACAACTAT	149						
QY	61	AGATGAGTGCTTAAACGAGGTCAAACAAAGTGTGTCATCATCCACACAGTGGCTTTGGG	120						
Db	150	AAATGGGTGCTTAAACGAGGTCAAACAAAGTGTGCTCATCCACACAGTGGCTTTGGG	209						
QY	121	GGCCTTCGACGCTCAAGAACTAGAGAGCTGTCCAAATGACACGACGTTTACACACATA	180						
Db	210	GGCCTTCGAGGCTCAAGAACTAGAGAGCTGTCCAAATGACACGAGGTTTACACACATA	269						
QY	181	TGCTTCAGATCAAGTCAACAAACATGGCCCTCATTTGCTTTTGGGGCCCTTTCATCAGG	240						
Db	270	TGCTTCGATCAAGTCAACAAACATGGCCCTCATTTGCTTTTGGGGCCCTTTCATCAGG	329						
QY	241	AAATGAGCTGTCTCTACAGGCTCCATCCAGCTTGAGAAATTAAGGATTA	289						
Db	330	AAATGAGCTGTCTCTACAGGCTCCATCCAGCTTGAGAAATTAAGGATTA	378						

[illegible]

```
FEATURES
  source      location/Qualifiers
1. 625
   /Organism="Homo sapiens "
   /db_xref="taxon:9606"
   /clone_lib="CT0261"
   /dev_stage="Adult"
```

```

172 a      154 c      150 g      149 t
      stringency conditions."
      mRNA and cDNA amplification were performed under low
      into the pUC18 vector. Reverse transcription of tissue
      716 - Ludwig Institute for Cancer Research) profiles
      from OHSR85 PCR (U. S. Letters Patent application No. 19
      Smal: A mini-library was made by cloning products derived
      /note="Organ: colon; Vector: puc18; Site_1: Smal; Site_2:

```

	Query Match	91.6%	Score 264.6	DB 10	length 625
	Best Local Similarity	98.2%	Pred. No. 1.1e-67		
	Matches 267	Conservative 0	Mismatches 5	Indels 0	Gaps 0
QY	18 GATCGAATATGTGCTGCTGACGATGGGGAAGACAACACTTAAGTGAGGCTTTAAG	77			
DB	61 GATCTGTAATTGTCTGCTGACGAGATGGGAAGACAACACTTAAGTGAGGCTTTAAG	120			
QY	78 AGGTCACAAAGTNGTGCCATATCCACACAGTGGCTTTGGGGCCCTTCGACGCTCAAG	137			
DB	131 AGGTAAACATGTGTGGTCCCTCATGACACAGTGGCTTTGGGGCCCTTCGACGCTCAAG	180			
QY	138 AACTAGAGAGAGCTGTCCAAATGACAGAGAGTTTACAGACATATGTTAGATCAAGTTC	197			
DB	181 AACTAGAGAGAGCTGTCCAAATGACAGAGAGTTTACAGACATATGTTAGATCAAGTTC	240			
QY	138 AGAACAATGGGCTCATTTGATGCTTTTGGGGCCCTTCATCAGAGAAATGAGCTGTCTTC	257			
DB	241 AGAACAATGGGCTCATTTGATGCTTTTGGGGCCCTTCATCAGAGAAATGAGAGCTGTCTTC	300			
QY	258 AGCGCTCATCCAGCTTGACAGTAAGGAGTTA	289			
DB	301 AGCGCTCATCCAGCTTGACAGTAAGGAGTTA	332			

LOCUS	621 bp	mRNA	EST	04-FEB-2000
AM361520				
QV2-CT0261-261099-011-e07				
DEFINITION				
AM361520				
ACCESSION				
AM361520.1				
GI:6866274				
KEYWORDS				
EST.				
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 621)			
AUTHORS	HCGP http://www.ludwig.org.br/ORSTES .			
TITLE	The FAPESP/LICR Human Cancer Genome Project			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Simpson A.J.G.			
	Laboratory of Cancer Genetics			
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,			
	Brazil			
	Tel: +55-11-2704922			
	Fax: +55-11-2707001			
	Email: asimpson@ludwig.org.br			
	This sequence was derived from the FAPESP/LICR Human Cancer Genome			
	Project. This entry can be seen in the following URL			
	(http://www.ludwig.org.br/scripts/gethtml2.pl?l1-QV2&l2-QV2-CT0261-			
	261099-011-e07&l3=1999-10-26&l4=1)			
	Seq primer: puc 18 forward			
	88			
	High quality sequence stop: 613.			
	High quality sequence stop: 613.			

```

FEATURES
  source
  {
    location/Qualifiers
    1..621
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_id="CR0261"
    /dev_stage="Adult"
    /note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:

```


polyA_signal	2912..2917	LRDHFNTSLQVNTTGTGLTPKRSSESEIEPEFELGATPFGNGDILFLAIDAVDSNLSKSEI SNLRKRSVEITPAQEPPTPEDSTPPCPDISTSINSTLIPGTHVLKIMKKNLGENGVYTLGLH
polyA_site	2933	/note="putative"
BASE COUNT	845 a 722 c 697 g 669 t	
ORIGIN		

Query Match	69.4%	Score 200.6	DB 12	Length 2933
Best Local Similarity	80.9%	Pred. No. 1	2e-48	
Matches 233	Conservative	0	Mismatches 55	Indels 0
				Gaps 0

QY	1	GAAATATCCAACTGATGTGATCTGAAATATGTGCTGTGACGGATGGGGAAGACAAACACTAT	60
Db	1225	GAAATATCCAACTGATGTGATCTGAAATATGTGCTGTGACCGATGGGGAAGACAAACCAT	1284
OY	61	AAATGGTGTCCTTTAAACGAGCTCAAAACAAATGTGTGCCATATCCACACAGTGTGCTTTGGG	120
Db	1285	TAGCAGCTGCTTTGACCTGTGTGTGTAACAGAGCGGGCCATCAATCCATACAGTGTGGCCCTGGG	1344
OY	121	GCCCTCTGAGCTCAGAAACTGAGAGAGCTGTCCAAATATGACAGAGAGGTTACAGACATA	180
Db	1345	ACCGGCTGCCGCTTAAAGAGCTTGAGCAGCTGTCCAAATGACAGAGAGGCTCTCAGACATA	1404
OY	181	TGCTTCAGATCAAGTTCAGAACAAATGGCCTCATTTATGTCTTTTGGGGCCCTTTCATCAGG	240
Db	1405	CTCTTCGAGTACAGTTCCAGAACAAATGTCCTTTGTATGCTTTCCGAGCACTCTCTCCAGG	1464
OY	241	AAATGACACTGTCTCTCAGCCGCTCATCATCAGTTGAGAGTAAGGATTT	288
Db	1465	AAATGCGGAGATGCTCTACGACTCATTCAGCTGGAGAGACAGGGAGAT	1512

RESULT	9
BFS80957	
LOCUS	BFS80957
DEFINITION	BFS80957 747 bp mRNA EST 12-DEC-2000 60110065991 NCI_CGAP_Co24 Mus musculus cDNA IMAGE:422390 5', mRNA sequence.
ACCESSION	BFS80957
VERSION	BFS80957.1 GI:11654669
KEYWORDS	EST.
SOURCE	house mouse. Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 747)
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D.
TITLE	
JOURNAL	
COMMENT	

FEATURES
SOURCE
1..747
Location/Qualifiers
High quality sequence stopt: 615
Plate: LLM9813 row: e column: 15
http://image.llnl.gov
found through the I.M.A.G.E. Consortium/LLNL at:
DNA Sequencing by: Incyte Genomics, Inc.
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
ISSUE Procurement: Jeffrey E. Green, M.D.
CNA Library Preparation: Life Technologies, Inc.

FEATURES	
Source	Location/Qualifiers
1. 747	
	/organism="Mus musculus"
	/strain="FVB/N"
	/db_xref="taxon:10090"
	/clone="IMAGE:4223990"
	/clone.lib="NCI CGAP Co24"
	/lab.host="DH10B (T1 phage-resistant)"
	/note="Organ: colon; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 kb. Constructed by Life

Technologies. Note: this is a NCBI_CGAP Library."				
BASE COUNT	203 a	198 c	187 g	159 t
ORIGIN				
Query: Match			67.88;	Score 196; DB 11; Length 747;
Best Local Similarity			80.48;	Pred. No. 2e-47;
Matches 229; Conservative			0; Mismatches 56; Indels 0; Gaps 0;	

195 GAAGTATCCCACTGATCGATCGTGAAGATTGTGCTGCTGACCCGATGGGGAGGACCAACACCAT 254

Db	255	TAGCAGCTGCTTTTGACCTGGTGAAACAGAGGGGGCCATCATCCATTAAAGTGGCCCTGGG	314
Oy	121	GGCCCTCAGAGTCAGAAACTAGAGAGCTGTCGCAAAATGACAGAGAGTTTACAGACATA	180
Db	315	ACCGGCTGCGCTTAAGAGCTTTGAGCAGCTGTCCAAAATGACAGAGAGCCCTGCAGACATA	374
Ox	151	TGCTTCAGATCAAGTTCAACAAATGGCCTCATTTGATGCTTTTGGGGGCCCTTTCATCAGG	240
Db	375	CTCTTCGATCAGGTTCCAGAACAAATGGTCTGTGATCTTTCGACAGACATCTCTCAGG	434
Oy	241	AAATGAGACTGTCTCTCAGCGGCTCCATCCAGTTTATAGAGTAAAGG	285
Db	415	AAATCGGCGATCGCTCAGCACTTCATCCATCCAGTTGAGAGCAGCAGGG	479

RESULT 10
BS965023
LOCUS BS965023 704 bp mRNA EST 12-JUN-2001
DEFINITION G0283166861 NC1_CGAF_CO24 Mus musculus cDNA IMAGE:4966281 5',
mRNA sequence.
ACCESSION BS965023
VERSION BS965023
KEYWORDS GI:14352660
SOURCE EST.
ORGANISM house mouse.
Mus musculus
Eutheria; Metazoa; Chordata; Craniata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 704)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

CNA Library Preparation: Life Technologies, Inc.
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: LLAM10995 row: g column: 18
 High quality sequence stop: 663.
 Location/Qualifiers
 1..704

```

BASE COUNT      179 a      174 c      189 g      162 t
ORIGIN
/organism="Homo sapiens"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4986281"
/clone_lib="NCI CGAP_Co24"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: colon; Vector: pMW-SPOrt6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. primer: 01lgo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

```


RESULT 12	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT	FEATURES
A0558859/c	439 bp	DNA	GSS	29-MAY-1999									
H5_2089_B1_B10_MR_CIT	Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2089 Col=19 Row=D, DNA sequence.												
A0558859	439 bp	DNA	GSS	29-MAY-1999									
A0558859.1	GI:4918591	GSS											
human.													
Homo sapiens													
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.													
1 (bases 1 to 439)													
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., a Hood,J.													
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome													
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)													
99380589													
Contact: Mahairas GG, Wallace JC, Hood L													
High Throughput Sequencing Center													
University of Washington													
401 Queen Anne Avenue North, Seattle, WA 98109, USA													
Tel: (206) 616-3618													
Fax: (206) 616-3887													
Email: jwallace@u.washington.edu													
Clones may be purchased from Research Genetics (info@resgen.com)													
BAC end Web Server: http://www.htsc.washington.edu													
Plate: 2089 row: D column: 19													
Seq primer: M13 Reverse													
Class: BAC ends													
High quality sequence stop: 439.													
Location/Qualifiers													
1..439													
/organism="Homo sapiens"													
/db_xref="taxon:9606"													
/clone="Plate=2089 Col=19 Row=D"													
/clone_lib="CIT Approved Human Genomic Sperm Library D"													
/sex="male"													
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"													
BASE COUNT	115 a	124 c	84 g	115 t	1 others								
ORIGIN													
Query Match	55.88;	Score 161.2;	DB: 13;	Length 439;									
Best Local Similarity	97.68;	Prod. No. 3.2e-37;											
Matches 163;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps									
0Y	1	GAATATTCACATGATGATGATGAAATTTGCTGCTGACGATGGGAAAGACACACTAT	60										
DB	395	GAATATTCACATGATGATGATGAAATTTGCTGCTGACGATGGGAAAGACACACTAT	336										
0Y	61	AAGTGGGTCTTTAAACGAGGTCAAAAGTNGTCCATCATCCACACAGTGGCTTTGGC	120										
DB	335	AAGTGGGTCTTTAAACG											

REFERENCE	TITLE	AUTHORS	JOURNAL	COMMENT
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	Contact: Robert Strausberg, Ph.D. Email: rgs@strs-remail.nih.gov
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	Tissue Procurement: ATCC
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	CDNA Library Preparation: Life Technologies, Inc.
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	DNA Sequencing by: Incyte Genomics, Inc.
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	plate: LLAM10366 row: d column: 08
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	High quality sequence stop: 685.
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	Location/Qualifiers
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	1. 831
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	/organism="Homo sapiens"
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	/db_xref="taxon:9606"
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	/clone="IMAGE:4500439"
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	/clone_1db="NIH_MGC_93"
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	/tissue_type="transitional cell papilloma, cell line"
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	/lab_host="DH10B (phage-resistant)"
1	(bases 1 to 831)	NIH-MGC	http://mgc.ncl.nih.gov/	/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-df primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC library."
BASE COUNT	246 a	174 c	192 g	219 t
ORIGIN				
Query Match	55.3%	Score 159.8;	DB 11;	Length 831;
Best Local Similarity	73.6%	Pred. No. 9.6e-37;		
Matches 203;	Conservative	0;	Mismatches 73;	Indels 0;
Gaps				
QY	14	GATGATGTTGAAATTTGTCCTGCTACAGGATGGGGAAGACAACTAATGAGTCTTT	73	
DB	277	GATGATGTTGAAATTTGTCCTGCTACAGGATGGGGAAGACAACTAATGAGTCTTT	336	
QY	74	AACGAGTCAAAACAAGNGTGCATATCCACACAGTCGCTTTGGGGCCCTCGACGT	133	
DB	337	GATGAGTCAAAACAAGNGTGCATATCCACACAGTCGCTTTGGGGCCCTCGACGT	396	
QY	134	CAAGAACTAGAGAGAGCTGCCAAATGACAGAGGTTTACAGACATATGCTTACATCAA	193	
DB	397	GAAGCAGTAATAGAGATGACCAATTAACAGACAGACATCTATTTTATGTTTCAGATGAA	456	
QY	134	GTTTCAGAACATGGCTTATGATGCTTTTGGGGCCCTTCATCAGAAATGAGAGTCTC	253	
DB	457	GCTCAGAACATGGCTTATGATGCTTTTGGGGCCCTTCATCAGAAATGAGAGTCTC	516	
QY	254	TCTCAGCGCTCCATCCAGCTTGAGAGTAAAGGATTA	289	
DB	517	TCTCAGCGCTCCATCCAGCTTGAGAGTAAAGGATTA	552	
RESULT	14			
LOCUS	A0557463/C			
DEFINITION	HS-2001 AL H11 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2081 Col=21 Row=0, DNA sequence.			
ACCESSION	A0557463	504 bp	DNA	GSS
VERSION	A0557463.1	GI:4917195		
KEYWORDS	GSS.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 504)			

AUTHORS Mahatras, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT Contact: Mahatras Gg, Wallace Jc, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 2081 row: 0 column: 21
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 504.
Location/Qualifiers
1. 504
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=2081 Col=21 Row=0"
/clone_id="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-coli DH10B"

BASE COUNT 141 a 119 c 102 g 137 t 5 others

ORIGIN

Query Match
Best Local Similarity 55.2%; Score 159.6; DB 13; Length 504;
Matches 162; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GAAATATCAACGATGATCTGAAATGTGCTGTGACGGATGGGAGACACACTAT 60
DB 338 GAAATATCAACGATGATCTGAAATGTGCTGTGACGGATGGGAGACACACTAT 279
QY 61 AAGGGGCTTTAAGGAGGCAACAAAGTNGTCCATCATCCACACAGTGGG 120
DB 278 AAGGGGCTTTAAGGAGGCAACAAAGTNGTCCATCATCCACACAGTGGG 219
QY 121 GCCCTTCACACTCAGAGACTAGAGAGCTGCCAAATGACAGAG 167
DB 218 GCCCTTCACACTCAGAGACTAGAGAGCTGCCAAATGACAGAG 172

RESULT 15
BG953500 275 bp mRNA EST 12-JUN-2001
LOCUS CMA-CT0629-230101-696-e07 CT0629 Homo sapiens cDNA, mRNA sequence.
DEFINITION CMA-CT0629-230101-696-e07 CT0629 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG953500
VERSION BG953500.1 GI:14371671
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 275)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Rui Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl-CM4-CT0629-230101-696-e07&cl=2001-01-23&cl=1>)
Seq primer: puc 18 forward
High quality sequence start: 30
High quality sequence stop: 234.
Location/Qualifiers
1. 275
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="CT0629"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site: 1; Smat: Site: 2; from ONESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 55 a 72 c 65 g 83 t

ORIGIN

Query Match
Best Local Similarity 41.2%; Score 119; DB 11; Length 275;
Matches 136; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 61 AAGGGGCTTTAAGGAGGCAACAAAGTNGTCCATCATCCACACAGTGGG 120
DB 142 AAGGGGCTTTAAGGAGGCAACAAAGTNGTCCATCATCCACACAGTGGG 103
QY 121 GCCCTTCACACTCAGAGACTAGAGAGCTGCCAAATGACAGAGCTTTACAGACA 180
DB 192 GCCCTTCACACTCAGAGACTAGAGAGCTGCCAAATGACAGAGCTTTACAGACA 44
QY 181 TGCTTCAGATCAAGTTCAGACAAATGCC 208
DB 43 AGCTTCAGATCAAGTTCAGACAAATGCC 16

Search completed: April 3, 2002, 20:27:27
Job time: 40210 sec

This Page Blank (uspto)

XX	PI	Nakanishi A, Morita S;
XX	DR	WPI: 2001-355935/37.
DR	P-PSDB; AAB737716.	
PT	XX	New antisense nucleotide, useful for treatment and prevention of
PT	XX	bronchial asthma and chronic obstructive pulmonary disease
PS	XX	Claim 3; Page 82-84; 104pp; Japanese.
SO	XX	Sequence 2742 BP: 833 A; 616 C; 623 G; 670 T; 0 other:
Query Match	99.7%; Score 288; DB 22; Length 2742;	
Best Local Similarity	99.7%; Pred. No. 4.7e-87;	
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
OY	1	GAATATTCACATGATGTGATCTGAATATGCTGCTGACGATGGGGAACACACTAT 60
Db	1194	gaataatccaactgagctggaatctgctgctgagagatgggaacaacactat 1253
OY	61	AAAGTGGTCTTTAAACGAGGTCAAACAAAGTGTGCCATCATCCACAGTGGCTTTGGG 120
Db	1254	aagtgggtgagctttaaagaggtcaacaagaatggtgccatccacacagtcgcttggg 1313
OY	121	GCCCTCTGACGCTCAAGACTAGAGAGCTGTGCCAAATGACAGAGGTTTACAGACATA 180
Db	1314	gccctctgagagctcaagaactagagagctgtccaaatgacagaggttcaagacata 1373
OY	181	TGCTTCAGATCAAGTTCAGAACATAGGGCTCATTTGATGCTTTGGGGCCCTTTCATCAGG 240
Db	1374	tgcttcagatcaagttcaagaacaatagtgctcatgtgttttggggcccttccatcagg 1433
OY	241	AAATGAGCTGTCTCTCAGCGCTCCATCCATCAGCTTGACAGTAAAGGACTTA 289
Db	1434	aaatgagctgtctctcagcgctccatccatccagcttgagagtaaggagta 1482
RESULT 2		
AAE81927		
ID	AAE81927	standard; cDNA; 2745 BP.
XX	AAE81927;	
AC	13-JUN-2001	(first entry)
XX		
DT		
XX		
DE		Human ICACC-1 nucleotide sequence.
XX		
KM		ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
KM		interleukin 9 induced calcium activated chloride channel; IL-9;
KM		calcium activated chloride channel; anti-allergic; anti-asthmatic;
KM		anti-inflammatory; immunomodulatory; cystic fibrosis;
KM		Inflammatory bowel disease; autoimmune disease; ss.
OS		
XX		Homo sapiens.
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..2745
FT	/tag=	a
FT	/product=	"ICACC-1"

FT		/note= "IL-9 induced calcium activated chloride channel"
XX	W09944620-A1.	
XX	10-SEP-1999.	
XX		
PF	03-MAR-1999; 99WO-US04703.	
PR		
XX	03-MAR-1998; 98US-0076815.	
XX		
FA	(MAGA-) MAGAININ PHARM INC.	
XX	Holroyd KJ, Levitt RC, Maloy WL, Louhed J, McLane M;	
PI	Nicotinoids NC, Zhou Y, Dong Q;	
DR	MPf; 1999-550979/46.	
XX	P-PsDB; AAB74824.	
PT	New nucleic acid encoding calcium activated chloride channel, used to	
XX	identity, e.g. specific modulators for treating atopic allergy -	
PS	Claim 1; Fig 4B; 75pp; English.	
CC	The present sequence encodes the human interleukin 9 (IL-9) induced	
CC	calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins	
CC	have anti-allergic, anti-asthmatic, anti-inflammatory and	
CC	immunomodulatory activities. Compounds (A) that downregulate ICACC are	
CC	used to alleviate asthma (or more generally atopic allergy), while those	
CC	(B) that upregulate ICACC are used to treat cystic fibrosis (CF) and	
CC	inflammatory bowel disease (IBD) (or other autoimmune diseases).	
CC	Measuring levels of ICACC-1 can be used in the diagnosis of asthma	
CC	(increased levels) or IBD (reduced levels), also for monitoring	
CC	treatment of these conditions. The ICACC proteins can be used:	
CC	(1) to raise specific antibodies (Ab); useful: (a) as immunoassay	
CC	reagents, and (b) as therapeutic (Ab); (ii) as fragments or mutants, as	
CC	(A) to disrupt bonding between ICACC and its ligands; and (iii) to	
CC	identify modulators and binding partners. ICACC polynucleotides can be	
CC	used to generate transgenic animals or recombinant cells, used to screen	
CC	for antagonists, also as a source of therapeutic antisense agents or	
CC	diagnostic probes (for quantifying mRNA expression, e.g. for	
CC	identification of modulators).	
SQ	Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;	
Query Match	99.7%; Score 288; DB 20; Length 2745;	
Best Local Similarity	99.7%; Pred. No. 4, 7e-87;	
Matches 288; Conservative	0; Mismatches 1; Indels 0; Gaps 0	
QY	I GAATATTCACACTGATGATCTGAATTGTGCTGTGAGCGATGGGAAACAACACTAT 60	
DB	1194 gaatattccaactgtatgagtctgaatttgcctgcgcgaagtgaggaaacaacact 1253	
QY	61 AAGTGGTGCTTTAACGAGCTCAACAAGTNGTGCATCATCCACACAGTGCCTTTGGG 120	
DB	1254 aagttgggtgctttaacgaagtcataaaaagttggtccatcatcacaacgctgctt199 1313	
QY	121 GCCCTGTGAGCTCAGAGACTAGAGAGCTGTGCCAAATGACAGAGAGTTTACAGACATA 180	
DB	1314 gcccttgtcaagctcaagaactagagagctgtlccaaaaatgacacggagtltaacagcata 1373	
QY	131 TGCTTCAGATCAAATTCAAGAAACATGGCCTCATTTGCTTTGGGCGCTTTCATCAGG 240	
DB	1374 tgcctcagatcaagctcagaacaatgcccattgatgtctttggggcccttcaccagg 1433	
QY	241 AAATGAGCTGCTCTCAGCGCTCCATCCAAGCTTGAGAGTAAGGAGTTA 289	
DB	1434 aaatgagctgtctctcagcgctccatccagcttgagagtaaggatla 1482	
RESULT	3	
ID	AAM46124 standard; CDNA: 2825 BP.	


```

XX AC AAH46124;
XX 11-SEP-2001 (first entry)
XX DE Human CLCA1 cDNA, SEQ ID NO:26.
XX
XX Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
XX expression inhibition; antisense therapy; gene therapy;
XX chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
XX ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 23..2767
XX FT /tag="a
XX FT /product="Human CLCA1"
XX FT /transl_except=" (pos:476..478, aa:Lys)
XX
XX WO200138530-A1.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-JP08232.
XX
XX 24-NOV-1999; 99JP-0333479.
XX PR 27-APR-2000; 2000JP-0127589.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Nakanishi A, Morita S;
XX
XX WPI: 2001-355935/37.
XX P-PSDB: AAB73716.
XX
XX New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease -
XX
XX Example 5; Page 92-94; 104pp; Japanese.
XX
XX The invention relates to an antisense nucleotide targeted to the mouse
XX Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX relates to an antibody specific for the Gob-5 protein, medical and
XX diagnostic compositions containing the antisense nucleotide or the
XX antibody, and methods and kits for screening for compounds which inhibit
XX the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX The antisense oligonucleotides and antibody are therefore useful for the
XX treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents human CLCA1 cDNA.
XX
XX Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;
XX
XX
XX Query Match 99.7%; Score 288; DB 22; Length 2825;
XX Best Local Similarity 99.7%; Pred. No. 4.7e-87;
XX Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 GAAATATCAACTGATGATCTGAAATTTGCTGCTGACGGATGGGAAGACACACTAT 60
XX |||||||
XX 1216 gaataatccaactgactgaaatctgctgctgacgagatg999aagaacaactat 1275
XX
XX 61 AAGTGGGTCTTTAAGCAGCTCAACAAAGTNGTGCATCACCACAGAGTGCCTTGGG 120
XX |||||||
XX 1276 aagtg99ggtcttaacgagtgcaacaaga9t9tgcacatccacacagtcgctt999 1335
XX
XX 121 GCCCTCTGACGCTCAAGAACTAGAGAGCTGTCCAAAATGACAGAGGTTTACAGACATA 180
XX |||||||
XX Db 1336 gccctctgacgctcaagaactagagagctgtccaaaatgacagaggttacaagacata 1395
XX
XX 181 TGCCTCAGATCAAGTTCGAACAATGGCTCTATTGATGCTTTTGGGCGCTTTCATCAGG 240
XX |||||||

```

```

XX Db 1396 tgcctcaagatccaagtlccaagaacaatgacctcaatgacttctt999cccttccacagg 1455
XX
XX 241 AATGAGCGCTGCTCTCAGACGCTCCATCCAGCTTGAGAGTAAGGAGTTA 289
XX |||||||
XX Db 1456 aatgagctgtctctccagcagctccatccaagcttga9gtaagagatta 1504
XX
XX
XX RESULT 4
XX ID AAH34879 standard; cDNA; 2854 BP.
XX
XX AAH34879;
XX
XX AAH34879;
XX
XX 03-SEP-2001 (first entry)
XX
XX Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; chromosome 1; ss.
XX
XX Homo sapiens.
XX
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
XX PR 03-NOV-1999; 99US-0165280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI: 2001-235357/24.
XX F-PSDB: AAG75474.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1; Page 3462-3463; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens (P), where
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing
XX inactive proteins or to supplement the patient's own production of P.
XX Additionally, N may be used to produce the colon cancer-associated Ps,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAB7789 represent sequences used in the exemplification of the
XX present invention.
XX
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;
XX
XX
XX Query Match 99.7%; Score 288; DB 22; Length 2854;
XX Best Local Similarity 99.7%; Pred. No. 4.7e-87;
XX Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 GAAATATCAACTGATGATCTGAAATTTGCTGCTGACGGATGGGAAGACACACTAT 60
XX |||||||
XX Db 1228 gaataatccaactgactgaaatctgctgctgacgagatg999aagaacaactat 1287
XX

```


XX AAH32943 to AAH37195 and AAG73514 to AAG77768 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated PS,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XQ Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;

Query Match	99.7%	Score 288;	DB 22;	length 2867;
Best Local Similarity	99.7%	Pred. No. 4.8e-87;		
Matches 288; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

OY	1	GAAATATCAACTGATGTAGTCGAATTGGTGGTGACGAGTATGGGGAAGACAACACTAT	60
Db	1232	gaataatccaactgctgatcgtcaaatctgtcgtctgcagcgatgggaaagaacaacctat	1291
OY	61	AAGTGGTGCTTTAAGCAGGTCAAACAAGTAGTGCATCATCCACACAGTGCCTTTGGG	120
Db	1292	aagtgtgtctttaaacgagtgtaaacaaaagtgtgcattcatccacaacgcttcgtttgg	1351
OY	121	GCCCCCTCAGCGTCAAGAACCTAGAGAGCGTGCAAAATGACAGGAGGTTTACAGCATA	180
Db	1352	gccctctgcagctcctaagaactaaggagagctgtcccataatgacagayagtttcagacata	1411
OY	181	TGCTTCAGATCAAGTTTCAGAACAAATGGCCTCATTTGATGCTTTTGGGGCCCTTTCATCAG	240
Db	1412	tgcctcacagataaagittcagaacaatlgcctcatgatgtctttggggccctttcatcagg	1471
OY	241	AAATGAGACTGTCTCTCAGCGCTCATCATCAGTTAGANTTAAGGGATTA	289
Db	1472	aaatggagctgtctctcaccgctccatccacgtctgagatlaagygatta	1520

	RESULT	7
AAH35019	ID	AAH35019 standard; cDNA: 3109 BP.
XX		
XX	AC	AAH35019;
XX		
DT	03-SEP-2001	(first entry)
XX		
DE	Human colon cancer antigen encoding cDNA SEQ ID NO:2101.	
XX		
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection;	
KW	colorectal carcinoma; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200122920-A2.	
XX		
PD	05-APR-2001.	
XX		
PF	28-SEP-2000; 2000WO-US26524.	
XX		
PR	29-SEP-1999; 99US-0157137.	
PR	03-NOV-1999; 99US-0163280.	
XX		

PA (HGMA-) HUMAN GENOME SCI INC.
XX
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX
DR WRI: 2001-235357/24.
ER P-RSDB: AAG75614.
XX
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
PT ¹⁵
PS Claim 1, Page 3587-3588; 9803pp. English.

AAH32943 to AAH37195 and AA673514 to AAG73788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patient's own production of P. Additionally, N may be used to produce the colon cancer-associated ps, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAH77789 represent sequences used in the exemplification of the present invention.

N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922.

Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match	99.7%	Score 288;	DB 22;	Length 3109;
Best Local Similarity	99.7%	Pred. No. 4.9e-87;		
Matches 288;	Conservative	0;	Mismatches 1;	Indels 0;
			Gaps	0

[illegible]

RESULT#	8
PAZ09840	
ID	AAZ09840 standard; cDNA; 3111 BP.
ZX	
AC	AAZ09840;
XX	
TT	26-NOV-1999 :(first entry)
XX	
DE	Human membrane spanning protein MSP-5 cDNA fragment 2.
ZX	
KW	Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;
KW	Neoplastic disorder; immunological disorder; reproductive disorder;
TW	MSP-5; ds.


```

XX OS Homo sapiens.
XX PN WO9946380-A2.
XX PD 16-SEP-1999.
XX PF 09-MAR-1999; 99WO-US05073.
XX PR 13-MAR-1998; 98US-0039064.
XX PA (INCYTE-) INCYTE PHARM INC.
XX PI Tang YT, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;
XX PI Guegler KJ, Kaser MR, Baughn MR, Shah P;
XX DR WPI; 1999-551409/46.
XX DR P-PSDB; AAY33298.
XX PT New human membrane spanning proteins used to, e.g. prevent and treat
XX PT neoplastic disorders -
XX PS Example 1; Page 80-81; 81pp; English.
XX CC This invention describes novel human membrane spanning proteins (MSPs),
XX CC and the polynucleotides encoding them. The products of the invention are
XX CC used to diagnose, prevent and treat neoplastic, immunological and
XX CC reproductive disorders. This sequence encodes a human membrane spanning
XX CC protein MSP-5 fragment.
XX SQ Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match          99.7%; Score 288; DB 20; Length 3111;
Best Local Similarity 99.7%; Pred. No. 4.9e-87;
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAAATATCCAACTGATGATGTAATGCTGCTGCTGAGATGGGAAAGACACACTAT 60
DB 1227 gaataatcaactatgctgctgaatgtgctgctgacgagtgaggaaacacactat 1286
QY 61 AAGTGGGCTCTTAAAGAGAGTCAAAAGTNGTGCATCATCCACACAGTGGCTTTGGG 120
DB 1287 aagtgaggctctttaaagagagtgcaaaaagtggtccatccacacagcttgcttggg 1346
QY 121 GCCCTCTGAGCTCAAGAACTAGAGAGCTGCCAAATGACAGAGGTTTACAGACATA 180
DB 1347 gccctctgagctcaagaaactagagagctgtccaaaatgacagagaggtttacagacata 1406
QY 181 TGCCTCAGATCAAGTTCAAGAACATGGCCTCATTTGATGCTTTGGGCCCTTTTCATCAGG 240
DB 1407 tgcctcagatcaagttcagaacatggcctcatgttgccttggggcccttccatcag 1466
QY 241 AATGAGACTGCTCTCTCAGCGCTGCATCCAGCTTAGAGTAGGATTA 289
DB 1467 aatgagactgctctctcagcgctgcattccatccagcttgagtaggagatta 1515

RESULT 9
AAI29502 ID AAI29502 standard; cDNA. 3311 BP.
XX AC AAI29502;
XX DT 12-OCT-2001 (first entry)
XX DE C902P determined cDNA sequence.
XX KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
XX KW immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX OS Homo sapiens.

```

```

PN WO200149716-A2.
XX PD 12-JUL-2001.
XX PF 29-DEC-2000; 2000WO-US35596.
XX PR 30-DEC-1999; 99US-0476296.
XX PR 10-JAN-2000; 2000US-0480321.
XX PR 15-FEB-2000; 2000US-0504629.
XX PR 06-MAR-2000; 2000US-0519444.
XX PR 19-MAY-2000; 2000US-0575251.
XX PR 29-JUN-2000; 2000US-0609448.
XX PR 28-AUG-2000; 2000US-0649811.
XX PA (CORI-) CORIXA CORP.
XX PI Xu U, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;
XX PI King GE, Wang T, Jlang Y;
XX DR WPI; 2001-441847/47.
XX PT Colon tumor associated proteins and nucleic acids useful for the
XX PT prevention, diagnosis and treatment of colonic cancer -
XX PS Claim 2; Page 425-426; 472pp; English.
XX CC The present invention describes colon tumour associated proteins (I) and
XX CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
XX CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
XX CC (II) may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate colon tumour associated protein (TCAP)
XX CC expression, such as colonic cancer. For example, (I) and (II) may be
XX CC used to treat disorders associated with decreased expression by
XX CC rectifying mutations or deletions in a patient's genome that affect the
XX CC activity of TCAPs by expressing inactive proteins or to supplement the
XX CC patients own production of them. Additionally, (II) may be used to
XX CC produce the TCAP proteins, by inserting the nucleic acids into a host
XX CC cell culturing the cell to express the protein. (II) and its
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC polymerase chain reaction (PCR) and hybridisation assays to detect and
XX CC quantitate the presence of similar nucleic acids in samples, and
XX CC therefore which patients may be in need of restorative therapy. (I) may
XX CC also be used as antigens in the production of antibodies against TCAPs
XX CC and in assays to identify modulators of TCAP expression and activity.
XX CC Anti-(I) antibodies and antagonists may also be used to down regulate
XX CC TCAP expression and activity. The anti-(I) antibodies may also be used
XX CC as diagnostic agents for detecting the presence of TCAPs in samples
XX CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
XX CC and AAI24494 to AAI24523 represent nucleotide and amino acid sequences
XX CC given in the exemplification of the present invention.
XX SQ Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;

Query Match          99.7%; Score 288; DB 22; Length 3311;
Best Local Similarity 99.7%; Pred. No. 5.1e-87;
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAAATATCCAACTGATGATGTAATGCTGCTGCTGAGATGGGAAAGACACACTAT 60
DB 1545 gaataatcaactatgctgctgaatgtgctgctgacgagtgaggaaacacactat 1604
QY 61 AAGTGGGCTCTTAAAGAGAGTCAAAAGTNGTGCATCATCCACAGTGGCTTTGGG 120
DB 1605 aagtgaggctctttaaagagagtgcaaaaagtggtccatccacacagcttgcttggg 1664
QY 121 GCCCTCTGAGCTCAAGAACTAGAGAGCTGCCAAATGACAGAGGTTTACAGACATA 180
DB 1605 gccctctgagctcaagaaactagagagctgtccaaaatgacagagaggtttacagacata 1724
QY 181 TGCCTCAGATCAAGTTCAAGAACATGGCCTCATTTGATGCTTTGGGCCCTTTTCATCAGG 240
DB 1725 tgcctcagatcaagttcagaacatggcctcatgttgccttggggcccttccatcag 1784

```


This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 02:59:04 ; Search time 22700.8 Seconds

(without alignments)
210.023 Million cell updates/sec

Title: US-09-049-696-7

Perfect score: 289
Sequence: 1 GAAATATCCACTGATGATGAT.....AGCTTGAGAGTAAGGATTA 289

Scoring table: IDENTITY_MUC
Gapop 10.0 , Capext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_scs:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rtd:*
33: em_htgo_hum:*
34: em_htg_inv:*
35: em_htg_rtd:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	288	99.7	2022	9	AK024970	AK024970 Homo sapi
2	288	99.7	2826	9	AF127036	AF127036 Homo sapi
3	288	99.7	3311	6	AX193489	AX193489 Sequence
4	288	99.7	3311	9	AF039400	AF039400 Homo sapi
5	200.6	69.4	2937	10	AB017156	AB017156 Mus muscu
6	198.4	68.7	3079	9	AF095584	AF095584 Sus scrof
7	164.4	56.9	35278	9	AF039401	AF039401 Homo sapi
8	164.4	56.9	113764	9	HSJ651E10	AL1322002 Human DNA
9	164.4	56.9	164891	2	AL1358950	AL1358950 Homo sapi
10	159.8	55.3	3204	9	AF127035	AF127035 Homo sapi
11	159.8	55.3	3321	9	AK000072	AK000072 Homo sapi
12	159.8	55.3	3265	6	AX092338	AX092338 Sequence
13	118.6	41.0	1895	9	AK000138	AK000138 Homo sapi
14	111	38.4	2820	4	AF001263	AF001263 Bos tauru
15	111	38.4	3388	4	AF001262	AF001262 Bos tauru
16	111	38.4	3317	4	AF001261	AF001261 Bos tauru
17	109.4	37.9	2765	10	AF115852	AF115852 Mus muscu
18	109.4	37.9	3058	10	AF108501	AF108501 Mus muscu
19	109.4	37.9	3137	10	BC008147	BC008147 Mus muscu
20	106.2	36.7	2984	4	BTU36445	U36445 Bos tauru
21	106.2	36.7	3022	10	AF047838	AF047838 Mus muscu
22	106.2	36.7	3471	10	AF052746	AF052746 Mus muscu
23	94.2	32.6	2970	9	AF127980	AF127980 Homo sapi
24	94.2	32.6	2832	9	AF043977	AF043977 Homo sapi
25	94.2	32.6	3604	9	AB026833	AB026833 Homo sapi
26	94.2	32.6	4077	6	AX054697	AX054697 Sequence
27	87	30.1	140718	2	AL356270	AL356270 Homo sapi
28	87	30.1	175591	2	AC068071	AC068071 Homo sapi
29	84.6	29.3	3415	9	AF043976	AF043976 Homo sapi
30	77.4	26.8	140718	2	AL356270	AL356270 Homo sapi
31	46.6	16.1	124096	2	AL354988	AL354988 Homo sapi
32	45.4	15.7	392	11	G61053	G61053 SHC-85269
33	35.6	12.3	163253	2	AL162575	AL162575 Homo sapi
34	35.6	12.3	179008	2	AC079602	AC079602 Homo sapi
35	35.6	12.3	187847	2	AC023155	AC023155 Homo sapi
36	35.6	12.3	199768	2	AL592490	AL592490 Homo sapi
37	35.6	12.3	201123	2	AL139410	AL139410 Homo sapi
38	35	12.1	69126	9	AL354752	AL354752 Human DNA
39	34.8	12.0	7017	14	AF296093	AF296093 Porcine t
40	34.8	12.0	116236	9	AL354877	AL354877 Human DNA
41	34.8	12.0	188823	2	AC016990	AC016990 Homo sapi
42	34.6	12.0	147327	2	AC078905	AC078905 Homo sapi
43	34.6	12.0	299550	1	AF001511	AF001511 Bacillus
44	34.2	11.8	222	8	AF009121	AF009121 Parasitau
45	34	11.8	72417	2	AC016988	AC016988 Homo sapi

ALIGNMENTS

RESULT	1	AK024970	2022 bp	29-SEP-2000
LOCUS	AK024970	Homo sapiens cDNA: FLJ21317 fls, clone COL02275, highly similar to		
DEFINITION	AF127036 Homo sapiens calcium-activated chloride channel protein 1			
ACCESSION	AK024970	AK024970.1 GI:10437397		
VERSION	AK024970.1	GI:10437397		
KEYWORDS	oligo capping; fls (full insert sequence).			
SOURCE	Homo sapiens colon cDNA to mRNA, clone lib:COL clone:COL02275.			
ORGANISM	Homo sapiens			
REFERENCE	1 (sites)			
AUTHORS	Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shihabara,T., Tanaka,T., Nakamura,Y., Isegai,T. and Sugano,S.			
TITLE	NED0 human cDNA sequencing project			
JOURNAL	Unpublished (2000)			
REFERENCE	2 (bases 1 to 2022)			

AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.

JOURNAL Direct Submission
Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdm@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT MEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'-6' end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES
source
1. 2022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="COL02275"
/tissue_type="colon"
/note="cloning vector pME18SFL3"
1. 2022
/note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA"

BASE COUNT 612 a 472 c 453 g 485 t

ORIGIN

Query Match 99.7%; Score 288; DB 9; Length 2022;
Best Local Similarity 99.7%; Pred. No. 6.3e-84;
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAAATATCCAACTGATGATGAAATTTGCTGCTGACGAGTGGGGAACACACTAT 60
|||||
DB 407 GAAATATCCAACTGATGATGAAATTTGCTGCTGACGAGTGGGGAACACACTAT 466
|||||

QY 61 AAGTGGGCTTTAAAGAGTGCACAAAGTNGGCCATTCACACAGTCGCTTGGG 120
|||||
DB 467 AAGTGGGCTTTAAAGAGTGCACAAAGTNGGCCATTCACACAGTCGCTTGGG 526
|||||

QY 121 GCCCTTCGACCTCAAGAACTAGAGAGCTCTCCAAAATGACAGAGGTTTACAGACATA 180
|||||
DB 527 GCCCTTCGACCTCAAGAACTAGAGAGCTCTCCAAAATGACAGAGGTTTACAGACATA 586
|||||

QY 181 TGCCTTCAGATCAAGTTTCAGAAACATGGCTTATGATGCTTTTGGGCCCTTTTATCAGG 240
|||||
DB 587 TGCCTTCAGATCAAGTTTCAGAAACATGGCTTATGATGCTTTTGGGCCCTTTTATCAGG 646
|||||

QY 241 AATGAGACCTGCTCTCAGCGCTCCATCCAGCTTGAAGAGGATTA 289
|||||
DB 647 AATGAGACCTGCTCTCAGCGCTCCATCCAGCTTGAAGAGGATTA 695
|||||

RESULT 2
AF127036 2826 bp mRNA PRI 10-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)
DEFINITION mRNA, complete cds.
ACCESSION AF127036
VERSION AF127036.1 GI:4585468
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 2826)
Agnel, M., Verma, T. and Culouscou, J. M.
Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea

JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792

REFERENCE 2 (bases 1 to 2826)
AUTHORS Agnel, M., and Culouscou, J. M.
JOURNAL Direct Submission
Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrières, Neuilly-Malmaison 92500, France

FEATURES
source
1. 2826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="small intestine; colon"
1. 2826
/gene="CaCC1"
5. 2749
/gene="CaCC1"
/note="CaCC1"
/note="bovine epithelial chloride channel homolog"
/product="calcium-activated chloride channel protein 1"
/protein_id="P42548.1"
/db_xref="GI:4585469"
/translation="MPPKSSVPLILHLLEGALSNLSLIQNNNGEIVAVDPNP
EDFTLIQIKDVTQASLYLFEATGRRTFKVNAIILPETWTKADYVPRKLETKNA
DVLVAESTPGNDPEYTEQMGNGCGERHILHPDIACKLAETGPQGRVHEWAH
LKWGVDEVDNDEKPYLSNGRQAVRCSAGITCTNVVKKCOGSCYTKRCEKTVGL
YKGECEVLOSROTERKASINFAQVDSIVIEFCEQHNKEAPRKQKCNLSRTEVI
RSEDEKRTPTPTOPNPPEFSLQIGORVCLVLSGSMATGNLRLNNGOLEFL
LORVELSGWMTFDSAAHVSELIQINSRDLRLARLPAAAGSGSISGLSEA
FYIRKKRYPDGSEIYLIDGENTISGCNEPKQSGALIIHYALPASAQLEELSK
MTGIDQTVASDOVQNGNGLIDAFGLSSNGVARSISLESKLTLSQNSOMNGTIV
DSVNGDQTLVLIWTPQPLIDPDSGQKGGVVDKNTKMAVLIQIPGIARGVTKY
SLQASQTLVLTYSRASNATLPIDVTSKNTGDSKFPSPLVYANIRQASPIIRA
SVTALLESVGVKVTLELDNGADATKDDGVSRFTYDNGRYSVKRALGVN
AARRVYPOOSGALYIPGEMENDEIQMNPRIKDDVOHKOVCSRSRSGSFVAS
DVNAPLIPDLFPPGQITDILKATIHGSLNLTWTAPGGDDYDNGTAKKYLIRISTILD
LRKFNESLDQVNTALIPKANSSEVPLFEPENITPENGDIPLIAQADKDKSEI
SNIAVSLFIPOTPEPSPDTSAPCPVIRHNSYTPGIIHLIKMKMIGELQSLIA

BASE COUNT 875 a 623 c 632 g 696 t

ORIGIN

Query Match 99.7%; Score 288; DB 9; Length 2826;
Best Local Similarity 99.7%; Pred. No. 6.7e-84;
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAAATATCCAACTGATGATGAAATTTGCTGCTGACGAGTGGGGAACACACTAT 60
|||||
DB 1198 GAAATATCCAACTGATGATGAAATTTGCTGCTGACGAGTGGGGAACACACTAT 1257
|||||

QY 61 AAGTGGGCTTTAAAGAGTGCACAAAGTNGGCCATTCACACAGTCGCTTGGG 120
|||||
DB 1258 AAGTGGGCTTTAAAGAGTGCACAAAGTNGGCCATTCACACAGTCGCTTGGG 1317
|||||

QY 121 GCCCTTCGACCTCAAGAACTAGAGAGCTCTCCAAAATGACAGAGGTTTACAGACATA 180
|||||
DB 1318 GCCCTTCGACCTCAAGAACTAGAGAGCTCTCCAAAATGACAGAGGTTTACAGACATA 1377
|||||

QY 181 TGCCTTCAGATCAAGTTTCAGAAACATGGCTTATGATGCTTTTGGGCCCTTTTATCAGG 240
|||||
DB 1378 TGCCTTCAGATCAAGTTTCAGAAACATGGCTTATGATGCTTTTGGGCCCTTTTATCAGG 1437
|||||

QY 241 AATGAGACCTGCTCTCAGCGCTCCATCCAGCTTGAAGAGGATTA 289
|||||
DB 1438 AATGAGACCTGCTCTCAGCGCTCCATCCAGCTTGAAGAGGATTA 1486
|||||

RESULT 3
AX193489 3311 bp DNA PAT 15-AUG-2001
LOCUS AX193489
DEFINITION Sequence 1056 from Patent WO0149716.
ACCESSION AX193489

VERSION AX193489.1 GI:15211440
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3311)
AUTHORS Xu, J., Lodes, M.J., Secrist, H., Benson, D.R., Meagher, M.J., Stolt, J.A., King, G.E., Wang, T., and Jiang, Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and methods for their use
JOURNAL Patent: WO 0149716-A 1056 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1. .3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 99.7%; Score 288; DB 6; Length 3311;
Best Local Similarity 99.7%; Pred. No. 6.8e-84;
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAATATCAACATGATGATCTGAATTTGCTGCTGACGGATGGGAGAACACACTAT 60
1545 GAATATCAACATGATGATCTGAATTTGCTGCTGACGGATGGGAGAACACACTAT 1604
DB 61 AAGTGGTGTCTTACGAGGTCAACAAAGTNGTCCATCATCCACACAGTCTTGGG 120
1605 AAGTGGTGTCTTACGAGGTCAACAAAGTNGTCCATCATCCACACAGTCTTGGG 1664
DB 121 GCCCTGACGCTCAAGACAGAGAGAGTGTCCAAATGACAGAGAGTTCAGACATA 180
1665 GCCCTGACGCTCAAGACAGAGAGTGTCCAAATGACAGAGAGTTCAGACATA 1724
QY 181 TGCCTCAGATCAAGTTCAGAGAGAGTGTCCATCATCATCTTTGGGCTTTCATCAGG 240
1725 TGCCTCAGATCAAGTTCAGAGAGAGTGTCCATCATCATCTTTGGGCTTTCATCAGG 1784
DB 241 AATGAGAGCTGTCTCTCAGAGAGTGTCCATCATCATCTTTGGGCTTTCATCAGG 289
1785 AATGAGAGCTGTCTCTCAGAGAGTGTCCATCATCATCTTTGGGCTTTCATCAGG 1833
DB

RESULT 4
LOCUS AF039400 3311 bp mRNA PRI 14-DEC-1998
DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA, complete cds.
ACCESSION AF039400
VERSION AF039400.1 GI:4009457
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3311)
AUTHORS Gruber, A.D., Eblie, R.C., Ji, H.L., Schreier, K.D., Fuller, C.M. and Pauli, B.U.
TITLE genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of Ca²⁺-activated Cl⁻ channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 3311)
AUTHORS Gruber, A.D., Eblie, R. and Pauli, B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES
source location/Qualifiers
1. .3311

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="p22-p31"
/tissue_type="small intestine"
1. .3311
/gene="hCLCA1"
352. .3096
/gene="hCLCA1"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPKSSVFLLILHLEGALSNSLIOLNNGYGVIAVDPNP
EDFLIOIDIMVYQASLILPEAFKGRFPERKNVAILIPMKPTADVVRKLEPKNA
DLVAESTPGNDEPTPEYOMKNGCEKERRIHLPDIFLAGKLIATYQGFAPFENAL
LRMGVEFEDYNNDEFEYLSNGRIQAVRCSAGITGVNVKCKOGGSCYTRKCTFNKVTGL
YEKCEFEVLOROTERKASIMEAHDVDSIVECEEDONNNKRNPNONKONKILRSTVEY
RDEDEFKTTPMTOPNPTEFSILOIGRIYCLVLDKSGSMATGNRLNLOAGOLF
LQVIELSGWGMVTFDSAAHVOSLQIOLNSGSDRLTAKRLPAAAGSTISCSGLRSLA
FTVIRKVPYDGSFVILITDGEDMTISGCPNENYQSGAILHTVALGSAOELBELSK
MTGSLQTYADQVONNCLIDAFGLSSGNKAVSORSIOLBSKGLTLONSGMNCTVIL
DSTVGKDTLTLITWTTPQPOLMDPSGOKGGVVDKNTKMAVLIQIPGAKVTWRY
SLAASQTLTLITVYSRASNATLPIITVTSINKDYSPEPLVYVANIROGASPIIRA
SVTALIESVNGKTYTLELDNGAGADATKDDGYRSYFTYDTRGYSVYKALGVNA
AARRVYIPDGLIPQIOTIDKAEIHGSLNLTWTAIPDDYDHRGKHYIIRISYILD
LRDFNLSLOVNTALIPKEANSEVLFRENTITFENGDTLFAIDAVDVLDKSEI
SNIAVSLFTIPQTPPETPSPDSAPCPNHNSTIPGHIILKIMRWIGELQLSIA

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 99.7%; Score 288; DB 9; Length 3311;
Best Local Similarity 99.7%; Pred. No. 6.8e-84;
Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAATATCAACATGATGATCTGAATTTGCTGCTGACGGATGGGAGAACACACTAT 60
1545 GAATATCAACATGATGATCTGAATTTGCTGCTGACGGATGGGAGAACACACTAT 1604
DB 61 AAGTGGTGTCTTACGAGGTCAACAAAGTNGTCCATCATCCACACAGTCTTGGG 120
1605 AAGTGGTGTCTTACGAGGTCAACAAAGTNGTCCATCATCCACACAGTCTTGGG 1664
DB 121 GCCCTGACGCTCAAGACAGAGAGTGTCCAAATGACAGAGAGTTCAGACATA 180
1665 GCCCTGACGCTCAAGACAGAGAGTGTCCAAATGACAGAGAGTTCAGACATA 1724
QY 181 TGCCTCAGATCAAGTTCAGAGAGAGTGTCCATCATCATCTTTGGGCTTTCATCAGG 240
1725 TGCCTCAGATCAAGTTCAGAGAGAGTGTCCATCATCATCTTTGGGCTTTCATCAGG 1784
DB 241 AATGAGAGCTGTCTCTCAGAGAGTGTCCATCATCATCTTTGGGCTTTCATCAGG 289
1785 AATGAGAGCTGTCTCTCAGAGAGTGTCCATCATCATCTTTGGGCTTTCATCAGG 1833
DB

RESULT 5
LOCUS AB017156 2937 bp mRNA ROD 10-NOV-1999
DEFINITION Mus musculus gob-5 mRNA, complete cds.
ACCESSION AB017156
VERSION AB017156.1 GI:3721911
KEYWORDS gob-5.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Komiyama, T., Tanigawa, Y. and Hirohashi, S.

TITLE
Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice

JOURNAL
Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)

REFERENCE
2 (bases 1 to 2937)

AUTHORS
Komiyama, T., Tanigawa, Y. and Hirohashi, S.

TITLE
Submitted (26-AUG-1998) to the DBJ/EMBL/Genbank databases. Tohru Komiyama, ERATO, JST, Genome Asymmetry Group, Doi Biosymmetry Project, Tokodai 5-9-9, Tsukuba, Ibaragi 300-2635, Japan (E-mail:tkom@doj.eriato.tokai.ac.jp, tel:01-298-48-1515, Fax:01-298-47-8901)

FEATURES
source
1. 2937
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="goblet cell"
/dev_stage="adult"
/tissue_type="intestine"
15. 2756
/gene="gob-5"
15. 2756
/gene="gob-5"
/codon_start=1
/protein_id="BAA33743.1"
/db_xref="GI:3721912"
/translation="MSELSKPYELLHLLEGVLSSESLIOLNNGYEGIVIAIDHDVP
EDFALIOHKKMVAQAPYLFATGKREYKNAVLIIPSMKAPETPKLETENNA
DYLVSTTSLGNDDEYTHIGAGEKGIHITLPELAKKLTOYGPORPTVHEMAH
FMGVFNEYNDKEFTLSKGRQAVHCSAATGKNVRCOGSGCTTNKCYIDRTG
LKNDVCEVDPDHQNKASIMFNQININSVEFTEHNQENADNOCNLSSTWEL
IOESDEKQTPMTAOPAPPTSLQIGRIYCLVLDKSGMLNDRLRMQASHLF
LQTVGSGWCVMTFEDSAAYVSELKOLNSGADRLIKHLPTVSAGTSGISGART
LAVTIVKKKPTGSEIVLTDGEDNTISCFDLKOSGALIHVALGPAAELEGLS
KMTGGLQTSNDOVONGLVDAFALISGNAIAHSLIOLESRGVLIQNNOMNGSVI
VNSVGKDTLFLITTTTHTPTPIFINDPSVEVNGKILDTTQAVAYGVGTAKVGRK
YSIQSSQTLTLTVTSRASATLPTITVPVKNKTGKPPSPYTVASIRQASPLR
ASVTALIESVNGKTVLELDNGAGADKNDGVSRFTAPDANGRYSVKIMALGV
TSDRORAPKRNAMTIDGIEDGEVRNMPREPSTSYODKOLCFSTSGSGSEIART
VPAAPITPDLPPCKITDLKASIOGONLVNLWTAPDGDYDHGRASNYIIRMSSTVD
LDRHFTSLQVNTTGLIPKESASEEIEFELGNTNGTDLFIATQVDSKLSKEI
SNIRSVFIPAPDEPPIPDSPPCDISINSTIPIGIVHLKIMMKLGMQVTLGLH"

BASE COUNT
860 a 718 c 693 g 666 t

Query Match
Best Local Similarity 69.4%; Score 200.6; DB 10; Length 2937;
Matches 233; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Db 1211 GAAATATCAACGATGATCTGAATTTGCTGCTGACGAGTGGGAAAGACAACACTAT 60
|||||
Db 1211 GAAATATCAACGATGATCTGAATTTGCTGCTGACGAGTGGGAAAGACAACACTAT 1270
|||||
Db 61 AAGTGGTCTTTTAAAGAGTCAACAAGTNGTCATATCCACAGAGTGGTGGG 120
|||||
Db 1271 TAGAGAGTCTTTTACCTGGTGAAGAGAGGCGGCATCATCATCAGTGGCTGGG 1330
|||||
Db 121 GGCCTCTGACGCTCAAGATAGAGAGCTGTCCAAATATGACAGAGTTTACAGACATA 180
|||||
Db 1331 ACCGGCTGCGCCGTAAGAGAGCTGTGAGAGCTGTCCAAATATGACAGAGCGCTGACACATA 1390
|||||
Db 181 TGCTTCAAGATCAAGTCAAGAAATGCGCTGATGATGCTTTTGGGCGCTTTATCAGG 240
|||||
Db 1391 CTCTTCGAGATGAGTTCAGAAACATGCTTGTGATGCTTTCACACACTCTCTCAGG 1450
|||||
Db 241 AAATGAGCTGTCTTCAAGAGCTTCATCCACTGAGAGTAAAGGATT 288
|||||
Db 1451 AAATGAGCTGTCTTCAAGAGCTTCATCCACTGAGAGTAAAGGAGT 1498
|||||

RESULT 6
AF095584

LOCUS
AF095584 3079 bp mRNA 13-OCT-2000

DEFINITION
Sus scrofa epithelial chloride channel protein (ABCC) mRNA,

ACCESSION
AF095584

VERSION
AF095584.1 GI:6002645

KEYWORDS
pig.

SOURCE
ORGANISM
Sus scrofa

REFERENCE
1 (bases 1 to 3079)
Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

AUTHORS
Gaspar, K.J., Racette, K.J., Gordon, J.R., Loewen, M.E. and Forsyth, G.W.

TITLE
Cloning a chloride conductance mediator from the apical membrane of porcine ileal enterocytes

JOURNAL
Physiol. Genomics (online) 3 (2), 101-111 (2000)

PubMed
20473747

REFERENCE
11015605

AUTHORS
2 (bases 1 to 3079)
Gaspar, K.J., Gabriel, S.E., Racette, K.J. and Forsyth, G.W.

TITLE
Direct Submission

JOURNAL
Submitted (28-SEP-1998) Veterinary Physiological Sciences,
University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4,
Canada

FEATURES
source
1. 3079
/organism="Sus scrofa"
/db_xref="taxon:9823"
/tissue_type="ileal mucosa"
1. 3079
/gene="ABCC"
133. 2886
/function="involved in exocrine chloride conductance"
/codon_start=1
/product="epithelial chloride channel protein"
/protein_id="AAF00077.1"
/db_xref="GI:6002646"
/translation="MSEFSSFLIVLHLEAGNSLSIOLNNGYEGIVIAIDNPV
EDERLQNKDVTASPYLFATGKREYKNAVLIIPSMKAPETPKLETENNA
DYLVSTTSLGNDDEYTHIGAGEKGIHITLPELAKKLTOYGPORPTVHEMAH
FMGVFNEYNDKEFTLSKGRQAVHCSAATGKNVRCOGSGCTTNKCYIDRTG
LKNDVCEVDPDHQNKASIMFNQININSVEFTEHNQENADNOCNLSSTWEL
IOESDEKQTPMTAOPAPPTSLQIGRIYCLVLDKSGMLNDRLRMQASHLF
LQTVGSGWCVMTFEDSAAYVSELKOLNSGADRLIKHLPTVSAGTSGISGART
LAVTIVKKKPTGSEIVLTDGEDNTISCFDLKOSGALIHVALGPAAELEGLS
KMTGGLQTSNDOVONGLVDAFALISGNAIAHSLIOLESRGVLIQNNOMNGSVI
VNSVGKDTLFLITTTTHTPTPIFINDPSVEVNGKILDTTQAVAYGVGTAKVGRK
YSIQSSQTLTLTVTSRASATLPTITVPVKNKTGKPPSPYTVASIRQASPLR
ASVTALIESVNGKTVLELDNGAGADKNDGVSRFTAPDANGRYSVKIMALGV
TSDRORAPKRNAMTIDGIEDGEVRNMPREPSTSYODKOLCFSTSGSGSEIART
VPAAPITPDLPPCKITDLKASIOGONLVNLWTAPDGDYDHGRASNYIIRMSSTVD
LDRHFTSLQVNTTGLIPKESASEEIEFELGNTNGTDLFIATQVDSKLSKEI
SNIRSVFIPAPDEPPIPDSPPCDISINSTIPIGIVHLKIMMKLGMQVTLGLH"

BASE COUNT
892 a 758 c 755 g 674 t

Query Match
Best Local Similarity 68.7%; Score 198.4; DB 4; Length 3079;
Matches 232; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Db 1326 GAAATATCAACGATGATCTGAATTTGCTGCTGACGAGTGGGAAAGACAACACTAT 1385
|||||
Db 1326 GAAATATCAACGATGATCTGAATTTGCTGCTGACGAGTGGGAAAGACAACACTAT 1385
|||||
Db 51 AAGTGGTCTTTTAAAGAGTCAACAAGTNGTCATATCCACAGAGTGGTGGG 120
|||||
Db 1386 AAGCGCTGCTTTCGAGAGTGAAGAGAGGAGGATATCATCATCAGTGGCTGG 1445
|||||
Db 121 GGCCTCTGACGCTCAAGATAGAGAGCTGTCCAAATATGACAGAGTTTACAGACATA 180
|||||

Db	1446	GCCCCGGCAGCAGGAAGAGCGTGGAGGAGCGTGTCCACATATGACAGCTGGTTTGCAGACATA	1505
Oy	181	TGCTTCAGATCAAGTTCAGAAACAATGGCCATCATGATGCTTTTGGGCCCCCTTTCATCAGG	240
Db	1506	TGCTTCAGACCAAGGCGCCGACCAACAGGCGCTTCATGAGCGTTTGGGGGCCCTTTCGTCGG	1565
Oy	241	AAATGGACCTCTCTCTGAGCGCTCCATCCAGCTTGACAGTGAAGGATTA	289
Db	1566	GACAGAGCTCTCTCCAGCGCTCCATCCAGCTGGAAGATCAGGCGTTA	1614
RESULT	7		
LOCUS	AF039401		
DEFINITION	Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene,		
ACCESSION	AF039401		
VERSION	AF039401.1		
KEYWORDS	GI:4009459		
SOURCE	human.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Gruber,A.D., Ehble,R.C., Ji,H.L., Schreur,K.D., Fuller,C.M. and Pauli,B.O.		
TITLE	Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of Ca2+-activated Cl- channel proteins		
JOURNAL	Genomics 54 (2), 200-214 (1998)		
MEDLINE	99047526		
REFERENCE	2 (bases 1 to 35278)		
AUTHORS	Gruber,A.D., Ehble,R. and Pauli,B.O.		
JOURNAL	Direct Submission		
FEATURES	Submitted (19-DEC-1997) Department of Pathology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA		
SOURCE	Location/Qualifiers		
	1..35278		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="p22-p31"		
	1..1617		
	/gene="hCLCA1"		
	1..33522		
	/gene="hCLCA1"		
	767..1265		
	/rpt_family="L1 transposable element"		
	/rpt_type="dispersed		
	1587..1596		
	/gene="hCLCA1"		
	/note="putative"		
	1618..1708		
	/gene="hCLCA1"		
	number=1		
	join(1618..1708,1962..2383,6667..6807,6979..7126,9656..9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33522)		
	/gene="hCLCA1"		
	/product="calcium-dependent chloride channel-1"		
	1962..2383		
	/gene="hCLCA1"		
	number=2		
	join(2222..2383,6667..6807,6979..7126,9656..9801,15454..15631,18591..18809,19783..20010,22253..22427,24628,26644..26859,27447..27708,28766,31840..32079,32919..33310)		
	/gene="hCLCA1"		
	/codon_start=1		
	/product="calcium-dependent chloride channel-1"		
	/protein_id="AAC95429.1"		
	/db_xref="GI:4009460"		
CDS			
exon			

[illegible]

Db 22384 GCCCTCGAGCTCAGAACTAGAGAGCGTGTCCAAATGACAGGTG 22430

RESULT 8
LOCUS HSJ651E10 113764 bp DNA PRI 27-MAY-2000
DEFINITION Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1, complete sequence.
ACCESSION AL122002
VERSION AL122002.16 GI:8247274
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 113764)

AUTHORS Brown, A.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT On Jun 4, 2000 this sequence version replaced gi:8247030.
requests: clonerequest@sanger.ac.uk
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Bm: EMBL; SW: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Ch1
RP4-651E10 is from the library RPCI-4 constructed at the Roswell Park Cancer Institute by the group of Piefer de Jong. For further details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP4-651E10.
Location/Qualifiers
1..113764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p22.3-31.1"
/clone="RP4-651E10"
/clone_1fb="RPCI-4"
BASE COUNT 34890 a 21989 c 22351 g 34534 t
ORIGIN

Query Match 56.9%; Score 164.4; DB 9; Length 113764;
Best Local Similarity 98.8%; Pred. No. 6.9e-43;
Matches 165; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAATATCCACATGATGTAATTCCTGCTGACGATGGGAAGACACACTAT 60
|||||
Db 21977 GAATATCCACATGATGTAATTCCTGCTGACGATGGGAAGACACACTAT 22036
|||||
QY 61 AAGTGGTGGTTTAAAGAGGTAACAAGAGTGGCATCATCCACACAGTGGTGG 120
|||||
Db 22037 AAGTGGTGGTTTAAAGAGGTAACAAGAGTGGCATCATCCACACAGTGGTGG 22096
|||||

QY 121 GCCCTCGAGCTCAGAACTAGAGAGCGTGTCCAAATGACAGAG 167
Db 22097 GCCCTCGAGCTCAGAACTAGAGAGCGTGTCCAAATGACAGGTG 22143

RESULT 9
LOCUS AL358950/c 164891 bp DNA HTG 23-JAN-2001
DEFINITION Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN PROGRESS ***. 10 unordered pieces.
ACCESSION AL358950
VERSION AL358950.4 GI:12539689
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 164891)

AUTHORS Plumb, B.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT On Jan 26, 2001 this sequence version replaced gi:3988471.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
Project Information
Center project name: BA444C12
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least Q40
Consensus quality: 163328 bases at least Q30
Insert size: 163991; sum-of-contigs
Insert size: 163373; 6.6% error; agarose-ff
Quality coverage: 5.19x in Q20 bases; sum-of-contigs quality
coverage: 5.21x in Q20 bases; agarose-ff

***** NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 34236: contig of 34236 bp in length
34237 34336: gap of 100 bp
34337 47372: contig of 13036 bp in length
47373 47472: gap of 100 bp
47473 67025: contig of 19533 bp in length
67026 67125: gap of 100 bp
67126 71239: contig of 4114 bp in length
71240 71339: gap of 100 bp
71340 83903: contig of 12564 bp in length
83904 84003: gap of 100 bp
84004 91357: contig of 7354 bp in length
91358 91457: gap of 100 bp
91458 96452: contig of 4995 bp in length
96453 96552: gap of 100 bp
96553 148189: contig of 51637 bp in length
148190 148289: gap of 100 bp
148290 159778: contig of 11469 bp in length
159779 159878: gap of 100 bp
159879 164891: contig of 5013 bp in length.
Location/Qualifiers
1..164891
/organism="Homo sapiens"

FEATURES
Source

LOCUS	AK000072	3221 bp	mRNA	PRI	22-FEB-2000
DEFINITION	Homo sapiens CDNA FLJ20065 fis, clone COL01613, highly similar to EC1C_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.				
ACCESSION	AK000072				
VERSION	AK000072.1	GI:7019922			
KEYWORDS	oligo capping; fis (full insert sequence).				
REMARKS	Homo sapiens colon CDNA to mRNA, clone_lib:COL clone:COL01613.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (sites) Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.				
TITLE	NEDO human cDNA sequencing project				
REFERENCE	Unpublished (2000)				
AUTHORS	2 (bases 1 to 3221) Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-FEB-2000) to the DDBJ/EMBL/Genbank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)				
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; CDNA full insert sequencing: Research Association for Biotechnology; CDNA library construction, 5'- < 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).				
FEATURES					
source	location/Qualifiers 1..3221 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="COL01613" /clone_lib="COL" /tissue_type="colon" /note="cloning vector pME18SFL3" 1..3221 /note="highly similar to EC1C_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"				
BASE COUNT	1105 a 600 c 634 g 882 t				
ORIGIN					
Query Match	55.3%; Score 159.8; DB 9; Length 3221; Best Local Similarity 73.6%; Pred. NO. 1.3e-41;				
Matches	203; Conservative 0; Mismatches 73; Indels 0; Gaps 0;				
0Y	14 GATGATCTGGAATTTGCTGCTGCTGACGAGTGGGGAAGACACATCTAAGTGGGTCCTT 73				
DB	1252 GATGATCTCGAAGTACTGCTGCTGACTGATGATGGGAGATTAACATGCAAGTCTGTAT 1311				
0Y	74 AACGAGGTCAACAAGAATGTCGCATCATCCACACAGTGGCTTTGGGGGCGCTCGAGCT 133				
DB	1312 GATGAGTCAACAAGAATGTCGCATCATCCACACAGTGGCTTTGGGGGCGCTCGAGCT 1371				
0Y	134 CAAGACTAGAGAGCTGTCAAATATGACAGAGGTTTACAGACATATGCTTCAGATCAA 193				
DB	1372 GAAGCAGTATATAGATGAGCAAGATTAACAGAGAGAAATCATTTTATGTTTCAGATGAA 1431				
0Y	194 GTTCGAGCAACAATGGCTCATTTGATGCTTTTGGGGCGCTTCATACGAATGAGAGCTGTC 253				
DB	1432 GCTCAGAAACAATGGCTCATTTGATGCTTTTGGGGCGCTTCATACAGAAATATCTGATCTC 1491				
0Y	254 TCTCAGCGCTCCATCCAGCTTGAGAGTAAGGATTA 289				
DB	1492 TCCCAAGATCCCTTCAGCTCGAAGATTAAGGATTA 1527				

AX092338	AX092338	3265 bp	DNA	PAT	21-MAR-2001
LOCUS	AX092338	3265 bp	DNA		
DEFINITION	Sequence 69 from Patent WO0116318.				
ACCESSION	AX092338				
VERSION	AX092338.1	GI:13444483			
KEYWORDS					
SOURCE	human:				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 3265) Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A., Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and Wood,W.I.				
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding the same				
JOURNAL	Patent: WO 0116318-A 69 08-MAR-2001;				
FEATURES	Location/Qualifiers				
source	1..3265				
BASE COUNT	1159 a 596 c 632 g 878 t				
ORIGIN					
Query Match	55.3%; Score 159.8; DB 6; Length 3265;				
Best Local Similarity	73.6%; Pred. No. 1.3e-41;				
Matches	203: Conservative 0; Mismatches 73; Indels 0; Gaps 0;				
QY	14 GATGATCTGGAATTTGTCTGCTGACGATGGGGAAGACACATATAGTGGTCTTT 73				
Db	1234 GATGATATCGGAAGTACTGCTGCTGACGTGAGGAGATACACATGCTTGTGATT 1293				
QY	74 AAGCAGGTCACAAAGTGTGTCATCATCCACACAGTCTGTGGGGCCCTCCAGCT 133				
Db	1294 GATGAAGTAAACAAAGTGGGCCATTTGTTATTTGCTTTGGGAAGAGCTGCTGAT 1353				
QY	134 CAAGAAGTACAGGAGAGCTGTCCAAATGACAGAGAGTTTACACACATATGCTTCAGATCAA 193				
Db	1354 GAGCAGATGATGAGATGACAAAGATTAACAGAGAGAGAAAGTATTTTATGCTTCAGATGAA 1413				
QY	194 GTTCAGACATAGGCGCTCATTTGATGCTTTTGGGCCCTTCATCCAGGAATGAGAGCTGC 253				
Db	1414 GCTCAGACAAATGCGCTCATTTGATGCTTTTGGGGCCTTACATCAGGAATATCTGATCTC 1473				
QY	234 TCTCAGCCCTCATCCAGCTTGAGAGTAAGGATTA 289				
Db	1474 TCCCAAGATGCCCTTCAGCTCGAAAGTAAGGATTA 1509				
RESULT 13					
LOCUS	AK000138	1895 bp	mRNA	PRI	22-FEB-2000
DEFINITION	Homo sapiens CDNA FLJ20131 fls, clone COL06357.				
ACCESSION	AK000138				
VERSION	AK000138.1	GI:7020030			
KEYWORDS	oligo capping; fls (full insert version).				
SOURCE	Homo sapiens colon CDNA to mRNA, clone_11b.COL clone:COL06357.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (sites) Kawabata,A., Hiki,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Ogaki,Y., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S. NEBO human cDNA sequencing project Unpublished (2000) 2 (bases 1 to 1895) Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio				

COMMENT

Sugano, Institute of Medical Science, University of Tokyo, Department of Virology; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cshai@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5' - 3' end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

Source

1. 1895
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL06357"
/clone_lib="COL"
/tissue_type="colon"
/note="cloning vector pME18SFL3"
66. 1475
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA90969.1"
/db_xref="GI:7020031"
/translation="MSKITGSHFYVDEAQNGLIDAFALITSGNDLSQKLSQLES
KGLTNSNANNDVVIDSTVGKDFELITNSLPPSISLDPGCTIMENKTVDATSK
MAYLSIGTAAGTAWAYLQAKAPETLITVSKAANSVPITVAKMNDVATSP
SPMIVAEILOGYVPLGANTAFLESGNHTLEVLNCGADSFNDCVYSRYFT
AYTNGRYSILKRAHAGANTAKLRLPRLNAAVLPGVNGEIEANPREIDDDQ
TTLEDPSRTAGGAFVYSQVSLPLPDQYPSQITDLDATYHEDKILITWAPGDNF
VGKQRIIRISASILDRLDSFDALQVNTDLSPKENSKESFAKFNENISENATH
IFIAIKSIDKSNLSKVSNIQVLTFLFOANPDIDPPTPTPDKSHNSGVNISTL
VLASGVVIVNFILSTPI"

CDS

BASE COUNT 658 a 352 c 357 g 528 t
ORIGIN

Query Match 41.0%; Score 118.6; DB 9; Length 1895;
Best Local Similarity 73.3%; Pred. No. 4.7e-28;
Matches 151; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 84 AACAAAGTNGGCCATCCACAGTCGCTTGGGGCCCTGACGTCAAGAACTAG 143
DB 1 AACAAAGTGGGGCCATTCATTTTATTTGTTGGGAAGCTGCTGATGAAGCATTA 60
QY 144 AGAGCTGTCCAAATGACAGAGGTTTACAGACATATCTTCAGATCAAGTTCAACA 203
DB 61 TAGAATGAGCAGATTAACAGAGAGATTTTATTTTTCAGATGAGCTCAGACAA 120
QY 204 ATGGCTCATTTGAGCTTTTGGGGCCCTTTCATCAGAGAAATGAGCTGTCTCAGCGCT 263
DB 121 ATGGCTCATTTGAGCTTTTGGGGCCCTTTCATCAGAGAAATGAGCTGTCTCAGCGCT 180
QY 264 CCATCCAGCTTGAGAGTAAGGATTA 289
DB 181 CCCTCAGCTCGAAGTAAGGATTA 206

RESULT 14
AF001263 2820 bp mRNA MAM 18-NOV-1997
LOCUS AF001263
DEFINITION Bos taurus clone 3 endothelial adhesion molecule Lu-ECAM-1 mRNA,
complete cds.
ACCESSION AF001263
VERSION AF001263.1 GI:2623766
KEYWORDS
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 2820)
AUTHORS Eblie,R.C., Widom,J., Gruber,A.D., Abdel-Ghany,M., Levine,R.,

Goodwin,A. and Pauli,B.U.
Cloning and characterization of Lu-ECAM-1 suggest it is an
endothelial chloride channel

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2820)
AUTHORS Eblie,R.C., Widom,J., Gruber,A.D., Abdel-Ghany,M., Levine,R.,
Goodwin,A. and Pauli,B.U.
TITLE Direct Submission
JOURNAL Submitted (25-Apr-1997) Vet. Pathology, Cornell U., VMC C4-146,
Ithaca, NY 14853, USA

FEATURES

Source

1. 2820
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="3"
/tissue_type="lung"
/cell_type="aortic endothelial cells"
194. 2656
/note="endothelial adhesion molecule; chloride channel"
/codon_start=1
/product="Lu-ECAM-1"
/protein_id="BAB86531.1"
/db_xref="GI:2623767"

CDS

BASE COUNT 940 a 565 c 560 g 755 t
ORIGIN

Query Match 38.4%; Score 111; DB 4; Length 2820;
Best Local Similarity 65.1%; Pred. No. 1.6e-25;
Matches 181; Conservative 0; Mismatches 91; Indels 6; Gaps 1;

QY 11 ACTGATGAGTCAAGTATGCTGCTGACGAGTGGGAGACACACTTAAGTGGCTGC 70
DB 1499 ACTTCTGTTCTGAATCACTATTACTGATGGGAGATATGAAATTAATTCATGC 1468
QY 71 TTTAAGAGGTCAACAAAGTNGGCCATTCATCCACAGTCGCTTGGGGCCCTTGCA 130
DB 1469 TTTGAGAGTGTAAAGCAAGTGTGCAATTCATCCACAGTCGCTTGGGGCCCTTGCT 1528
QY 131 GCTCAAGAACTAGAGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTCAGAT 190
DB 1529 GCCAAGAACTGAGACACTTGTCAATATGACAGAGGATATC-----GTTTTTGGCC 1582
QY 131 CAAGTTCAGAAAGTGGCTCATGATGCTTTGGGGCCCTTTCAGACAGAAATGAGCT 250
DB 1533 AATAAGACATTAACCTGCTTACTAATGCTTTCAGATGAGATTCATCTAAGAGTGAAC 1642
QY 251 GTCTCTCAGCCGCTCCATCCAGCTTGAAGTAAAGGATTT 288
DB 1613 ATCAGCTCAGCAGGCTATTCAGTTGGAACCAAGCCTT 1680

RESULT 15

AF001262 3288 bp mRNA MAM 18-NOV-1997
LOCUS AF001262
DEFINITION Bos taurus clone 2 endothelial adhesion molecule Lu-ECAM-1 mRNA,
complete cds.
ACCESSION AF001262
AUTHORS


```

VERSION      AF001262.1  GI:2623764
KEYWORDS
SOURCE       cow.
ORGANISM     Bos taurus
             Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
             Bovidae; Bovine; Bos.
REFERENCE    1 (bases 1 to 3288)
AUTHORS      Elble, R.C., Widom, J., Gruber, A.D., Abdel-Chany, M., Levine, R.,
             Goodwin, A. and Pauli, B.U.
TITLE        Cloning and characterization of Lu-ECAM-1 suggest it is an
             endothelial chloride channel
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 3288)
AUTHORS      Elble, R.C., Widom, J., Gruber, A.D., Abdel-Chany, M., Levine, R.,
             Goodwin, A. and Pauli, B.U.
TITLE        Direct Submission
JOURNAL      Submitted (25-APR-1997) Vet. Pathology, Cornell U., VMC C4-146,
             Ithaca, NY 14853, USA
FEATURES
SOURCE       Location/Qualifiers
             1..3288
             /organism="Bos taurus"
             /db_xref="taxon:9913"
             /cell_type="aortic endothelial cells"
             /clone="2"
             /tissue_type="lung"
             59..2443
             /note="endothelial adhesion molecule; chloride channel"
CDS
             /product="Lu-ECAM-1"
             /protein_id="AAB86530.1"
             /db_xref="GI:2623765"
             /translation="MVLCLNVYLFLLPLPGKMSVNLINNGYDGIIVAINPSVE
             DERIKENIKEMVPEASTYLPFATRRVYFRNVLIPMTKSKSEYFIPKESYDQAD
             VIANPYLTKGDDPYTLQYGRGCGEKGIHTFPFLTLNNEHYIGSRGVEVHEMHL
             RMGTFDEYNVDQPYISRKNTTEATRCSTHITGINVFKCPGSCITSLCRDSOTG
             LYEAKEFTFLPKSQJAKESIMFPLSHSVETCEKTHNTAPNLQKNCCKSTWDY
             IMNGVFOENTSPTMEMPPIHPFSLKSKRVVCLVLDKSGSMADRLQMNQAAE
             LYLQVTEKSGSLVGMTEDSVAEIOMHLTRITDENVYOKITAKLPQVANGSTICRGL
             KAGROATIHSDOSTSGSEIILLTDGEDNEINSCEDYKRSQAIHTITALGPSAKELE
             TLSMTGGRPREFAKDKITGLTNASRSISRSSTQAOIOLESKALKITGRKRYNGTV
             PLSVGVNDTFEVYWTIOKEIYLQDPKCKKRYTSPREKMLNRSARLQIPGLAFT
             GTWTVSLNNHNASOMLVTVTTNARSPTIPPVATAHMSOHTAHTSPMLVYQVSG
             GFPLVLGISVTAIIETEDGHQVLTLELDNCGARPTVKNDSIYSRYFTDYGENRYSLK
             VHAQRNNTARLNRPOPKNKYLVPYGVYENKILINPPEVKNKDLAKAKIEDFSRLT
             SGSEFTVGAPPPGNHPSVFPFSKTTDLFAFKEDYIQLSWTAPGNVLDKXES"
BASE COUNT  1043 a      707 c      645 g      893 t
ORIGIN
Query Match      38.4%; Score 111; DB 4; Length 3288;
Best Local Similarity 65.1%; Pred. No. 1.6e-25;
Matches 181; Conservative 0; Mismatches 91; Indels 6; Gaps 1;
0Y      11 ACTGATGATCTGAATTGCTGCTGACGCGATGGGGAAGACAACATTAAGTGGGTGC 70
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1274 ACTTCTGTTCTGAATCTACTATTAACTGATGGGGAAGATAAATAATTCAATGC 1333
0Y      71 TTTAAGAGAGTCAACAAGATNGTGCATCATCCACAGTCGCTTTGGGGCCCTCTGCA 130
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1334 TTTCAGAGATGTAAAGAGAGTGTGATCATCCACACCATTTGCTGTGGACCCCTGTGCT 1393
0Y      131 GCTCAAGACTAGAGAGCTGTCCAAATGACAGAGAGTTTACAGACATATGCTTCAGAT 190
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1394 GCCAANAGACTGGACACTTGTCAATATGACAGAGAGATATC-----GTTTITTTTGGC 1447
0Y      191 CAAGTTCAACAACATGGCCTCATGTGCTTTGGGGCCCTTTCATCAGAAATGGAGCT 250
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1448 AATAAGACATTAAGTGGCCTTACTAATGCTTTCAGTAGAATTCATCTAGAAAGTGAAGC 1507
0Y      251 GTCTCTACGCGCTCCATCCAGCTTGAGAGTAAGGATT 288
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1508 ATCACTCAGCAGGCTATTTCAGTTGGAAGCAAGACCTT 1545

```

Search completed: April 4, 2002, 02:59:52
 Job time: 60274 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:27 ; Search time 16681 Seconds

(without alignments)
162.980 Million cell updates/sec

Title: US-09-049-696-8

Perfect score: 253
Sequence: 1 AACAAAGTGTGCGATCATTG.....GGCAGAGTGTGTCGACAG 253

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estlov:*
9: em_hic:*
10: qb_est1:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_fod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249.8	98.7	654	10	AM361522 QV2-CT026
2	249.8	98.7	654	10	AM361521 QV2-CT026
3	248.2	98.1	625	10	AM361532 QV2-CT026
4	238.8	94.4	621	10	AM361520 QV2-CT026
5	238.8	94.4	652	10	AM753449 QV2-CT026
6	225.2	89.0	561	10	AM361523 QV2-CT026
7	174.6	69.0	661	11	BF581041 QV2-CT026
8	174.6	69.0	2933	12	AK008659 Mus muscu
9	162	55.0	2915	12	AK007466 Mus muscu
10	139.2	55.0	747	11	BF580957 602100659
11	130.8	51.7	831	11	BG286204 602383026
12	126.4	50.0	704	11	BG965023 602831668

104	41.1	550	11	BF355533 RC3-HT086
104	41.1	906	11	BF581342 RC3-HT086
101	39.9	534	11	BF355513 RC3-HT086
97	38.3	275	11	BG953500 CM4-CT062
96.4	38.1	455	10	AA871197 VQ32G11.T
93.4	36.9	858	11	BG675746 602621972
93.4	36.9	940	11	BG678187 602621972
85.2	33.7	510	10	AM762590 ur64b08.Y
82	32.4	1033	11	BG243305 602355738
81	32.0	462	10	AI851265 UI-M-BH0-
80.8	31.9	435	10	AA443258 aa14d09.T
77.6	30.7	439	13	AO558859 HS-2089.B
73.6	29.1	504	13	AO557463 HS-2081.A
72.6	28.7	468	10	AA726662 vu93f05.T
70.8	28.0	935	11	BF382180 601271754
69.4	27.4	584	11	BF578919 602095492
68	26.9	453	13	BF116637 uy99g07.Y
65.2	25.8	308	10	AO207683 HS-2249.A
65	25.7	866	11	AO707718 AV077718
59.6	23.6	314	10	BI155803 602904376
52.2	20.6	331	10	BB225759 BB225759
52	20.6	299	10	BI112685 602898565
51.4	20.3	517	12	AZ708112 RPCI-23-2
49.8	19.7	430	10	AF071903 AF071903
49	19.4	337	10	AZ706870 RPCI-23-2
41.8	16.5	285	10	BS560274 BS560274
39.8	15.7	633	10	BB499841 BB499841
				BI124752 BI124752
				AK006921 Mus muscu
				AI493356 tg70f11.x
				AV047614 AV047614
				BB226010 BB226010
				AM753451 QV2-CT026

ALIGNMENTS

RESULT 1
LOCUS AM361522 654 bp mRNA EST 04-FEB-2000
DEFINITION QV2-CT0261-261099-011-f10 CT0261 Homo sapiens CDNA, mRNA sequence.
ACCESSION AM361522
VERSION AM361522.1 GI:6866276
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 654)
HCGP <http://www.ludwig.org.br/ORESRES>.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL:
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV2&ct-QV2-CT0261-261099-011-f10&ct3-1999-10-26&ct4-1>)
Seq primer: puc 18 forward
High quality sequence start: 132
High quality sequence stop: 643.
Location/Qualifiers
1..654
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"


```

/dev:stage="Adult"
/node:"Organ: colon; Vector: puc18; Site:1: Sma1; Site:2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      173 a      156 c      162 g      163 t
ORIGIN

```

	Query Match	Best Local Similarity	Matches	251: Conservative	98.7%: 99.2%:	Score 249.8:	DB 10:	Length 654:	Pred. 1.3e-64:	Mismatches 2:	Indels 0:	Gaps 0:
OY	1	AACAAGTGGTCCATMCATCCACAGTGCCTTTGGGCCCTCTGCAGTCAAGAATAG	60									
DB	177	AACATAGTGTGTCCTCTCATCCACAGTGCCTTTGGGGCCCTCTGCAGTCAAGAATAG	236									
OY	61	AGGAGCTGTCCAAATGACAGGAGGTTTACAGCATATGCTTCAGATCAAGTTCAGAACA	120									
DB	237	AGGAGCTGTCCAAATGACAGGAGGTTTACAGCATATGCTTCAGATCAAGTTCAGAACA	296									
OY	121	ATGGCCTCATTTATGCTTTTGGGCCCTTTCAATCAGAGAAATGGAGCTGTCTTCAGCGCT	180									
DB	297	ATGGCCTCATTTATGCTTTTGGGGCCCTTTCAATCAGAGAAATGGAGCTGTCTTCAGCGCT	356									
OY	181	CCATCCAGCTTAAAGGATTAAGCCCTCAGACAGCCAGTGGATGAATGGACAG	240									
DB	357	CCATCCAGCTTAAAGGATTAAGCCCTCAGACAGCCAGTGGATGAATGGACAG	416									
OY	241	TGATGTGGACAG	253									
DB	417	TGATGTGGACAG	429									

[illegible]

```
FEATURES      location/Qualifiers
SOURCE        1. 657
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="CR0261"
              /dev_stage="Adult"
```

/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
 Sma1; A mini-library was made by cloning products derived
 from ORESPEs PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 175 a 160 c 164 g 158 t
 ORIGIN

	Query: March	Similarity	98.7%	Score 249.8	DB 10	Length 657
	Best Local	Similarity	99.2%	Pred. No. 1.3e-64		
	March 251	Conservative	0	Mismatches	2	Indels 0; Gaps 0;
QY	1	AACAAAGTGTCACATCCACACAGTCGCTTTTGGGCCCTTCAGCTCAAGAACTAG	60			
DB	123	AACAAAGTGTCGCTTCATCCACACAGTCGCTTTTGGGCCCTTCAGCTCAAGAACTAG	232			
QY	61	AGGAGCTGTCCAAATGACAGGAGGTTTACAGACATATGCTTCAGATCAAGTTCAACA	120			
DB	233	AGGAGCTGTCCAAATGACAGGAGGTTTACAGACATATGCTTCAGATCAAGTTCAACA	292			
QY	121	ATGGCCCTATTGATCTTTTGGGCCCTTATTCAGGAATGAGACTGCTCTAGCGCT	180			
DB	293	ATGGCCCTATTGATCTTTTGGGCCCTTATTCAGGAATGAGACTGCTCTAGCGCT	352			
QY	181	CCATCCAGCTTACAGACTAAGGATTAACCCCTCACAACAGCAGTGGATGAATGGCACAG	240			
DB	353	CCATCCAGCTTACAGACTAAGGATTAACCCCTCACAACAGCAGTGGATGAATGGCACAG	412			
QY	241	TGATCGTGGACAG	253			
DB	413	TGATCGTGGACAG	425			

FEATURES	REFERENCE	DEFINITION	LOCUS	RESULT
source	ORGANISM	VERSION	AM361532	3
location/Qualifiers	human.	KEYWORDS	AM361532.1	
	EST.		GI:6666286	
	Homio sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
	1 (bases 1 to 625)			
	HCGP http://www.ludwig.org.br/ORESTES .			
	The FAPESP/LICR Human Cancer Genome Project			
	Unpublished (1999)			
	Contact: Simpson A.J.G.			
	Laboratory of Cancer Genetics			
	Ludwig Institute for Cancer Research			
	Brazil			
	Brazil			
	Tel: +55-11-2704922			
	Fax: +55-11-2707001			
	Email: asimpson@ludwig.org.br			
	This sequence was derived from the FAPESP/LICR Human Cancer Genome			
	Project. This entry can be seen in the following URL			
	http://www.ludwig.org.br/scripts/gethtml2.pl?tl=GV2&lt2=GV2-CT0261			
	261099-011-e03<3=1999-10-26<4=1)			
	Seq primer: puc 18 forward			
	High quality sequence start: 93			
	High quality sequence stop: 608.			

```
source
1..625
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
```


SmaI: A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions.

BASE COUNT 172 a 154 c 150 g 149 t

ORIGIN

Query Match 98.1%; Score 248.2; DB 10; Length 625;
Best Local Similarity 98.8%; Pred. No. 3.9e-64;
Matches 250; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AACAAAGTGGTCCATCATCCACAGAGTGGGGCCCTCTGAGTCAAGACTAG 60
Db 127 AACATAGTGGTCCCTCATGCACAGTGGGGCCCTCTGAGTCAAGACTAG 186
QY 61 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAAGACA 120
Db 187 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAAGACA 246
QY 121 AAGGCTCATGATGCTTTGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCCGT 180
Db 247 AAGGCTCATGATGCTTTGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCCGT 306
QY 181 CCATCCAGCTTGAGAGTAAGGATTACCTCCAGAACAGCCAGTGAATGACAG 240
Db 307 CCATCCAGCTTGAGAGTAAGGATTACCTCCAGAACAGCCAGTGAATGACAG 366
QY 241 GTGATCGTGACAG 253
Db 367 GTGATCGTGACAG 379

RESULT 4
LOCUS AM361520 621 bp mRNA EST 04-FEB-2000
DEFINITION QV2-CT0261-261099-011-e07 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM361520
VERSION AM361520.1 GI:6866274
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 621)
AUTHORS HGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV2&tl2-QV2-CT0261-261099-011-e07&tl3=1999-10-26&tl4=1)
Seq primer: puc 18 forward
High quality sequence start: 88
Location/Qualifiers
1..621
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0261"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions.

BASE COUNT 166 a 151 c 150 g 154 t

ORIGIN

Query Match 94.4%; Score 238.8; DB 10; Length 621;
Best Local Similarity 98.8%; Pred. No. 2.6e-61;
Matches 251; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 AACAAAGTGGTCCATCATCCACAGAGTGGGGCCCTCTGAGTCAAGACTAG 60
Db 124 AACATAGTGGTCCCTCATGCACAGTGGGGCCCTCTGAGTCAAGACTAG 183
QY 61 AG-GAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAAGACA 119
Db 184 AGTGAAGTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAAGACA 243
QY 120 AAGGCTCATGATGCTTTGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCCGT 179
Db 244 AAGGCTCATGATGCTTTGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCCGT 303
QY 180 CCATCCAGCTTGAGAGTAAGGATTACCTCCAGAACAGCCAGTGAATGACAG 239
Db 304 CCATCCAGCTTGAGAGTAAGGATTACCTCCAGAACAGCCAGTGAATGACAG 363
QY 240 GTGATCGTGACAG 253
Db 364 GTGATCGTGACAG 377

RESULT 5
LOCUS AM753449 652 bp mRNA EST 28-APR-2000
DEFINITION QV2-CT0261-261099-011-e08 CT0261 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM753449
VERSION AM753449.1 GI:768381
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 652)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.G.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
COMMENT Contact: Simpson A.J.G.
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV2&tl2-QV2-CT0261-261099-011-e08&tl3=1999-10-26&tl4=1)
Seq primer: puc 18 forward
High quality sequence stop: 14.
Location/Qualifiers
1..652
/organism="Homo sapiens"


```

/db_xref="taxon:3606"
/clone_lib="CT0261"
/dev_strage="Adult"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT      179 a      153 c      161 g      159 t
ORIGIN

```

Query Match	94.4%	Score 238.8	DB 10	Length 652
Best Local Similarity	98.8%	Pred. No. 2.6e-61		
Matches 251	Conservative	0	Mismatches 2	Indels 1
				Gaps 1

0Y 1 AACAAAGTGGTCATCATCCACACAGTCGTTGGGGCCCTGCG -AGCTCAAGAACTA 59
 #####
 184 AACATAGTGGTGGCCCTCATCCACACAGTCGCTTTGGGGCCCTTGCAAGCTCAAGAACTA 243
 #####

07 GAGGAGGCTGTCCAAAATGACAGGAGGTTTACAGACATATGCTTCAGATCAGGTTTCAGAAC
 119
 244 GAGGAGGCTGTCCAAAATGACAGGAGGTTTACAGACATATGCTTCAGATCAGGTTTCAGAAC 303

120 AATGGCCCTCATGATGCTTTGGGGCCCTTTCATCAGGAATATGAGAGCTCTCAGGGC 175
 304 AATGGCCCTCATGATGCTTTGGGGCCCTTTCATCAGGAATATGAGAGCTCTCAGGGC 363

Db 364 TCCATCCAGCTTGAGAGTAAGGATTAAACCCCTCAGAACAGCCAGTGGATGAATGGCACA 423

Db 424 GTGATCGTGACAG 437

LOCUS	AM361523	561 bp	mRNA	EST	04-FEB-2000
RESULT	6				

AM361523 GI:6866277
 AM361523.1
 EST.
 KEYWORDS

ORGANISM	Homio sapiens	Chordata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa;		
Mammalia; Eutheria;		
Primates;		
Catarrhini; Homnidae; Homo.		

AUTHOR(S)
 HCGP <http://www.ludwig.org.dir/ONESTS>.
 TITLE
 The FAPESP/LICR Human Cancer Genome Project
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Simpson J I C

Laboratory Of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: 214-2707001
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/ICTR Human Cancer Genome

(<http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV2&t2=QV2-CT0261-261099-011-g05&t3=1999-10-26&t4=1>)
 seq primer: puc 18 forward

```

FEATURES
source      High quality sequence stop: 560.
            Location/Qualifiers
            1..561

```

/db_xref="taxon:9606"

```

/clone_lib="CT0261"
/dex_stage="Adult"
/notes="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
Sma1; A mini-library was made by cloning products derived
from ORESFES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      154 a      140 c      140 g      127 t
ORIGIN

```

Query Match	89.0%;	Score 225.2;	DB 10;	Length 561;
Best Local Similarity	98.0%;	Pred. No. 3e-57;		
Matches 249;	Conservative 0;	Mismatches 3;	Indels 2;	Gaps 2;

Db 80 ACATAGTGGTGGCCCTCATGCACACAGTGGCTTTGGGGGCCCTCGCAGCTCAAGAACTAG 139

140 AGTGAAGTGTCCAAAATGACAGAGAGTTTACAGACATATGTCTTCAGATCCAAAGTTACAGAAC 199

Dv	160	TCCATCCAGCTGAGACGTAAAGGGATTAAACCCCTTCAGAACCAAGCGTGGATGGATGGGACA	249
Db	200	AATGGCCCATTTGATGCTTTTTGGGGCCCCCTTTCATCAGGAAAAATGGAGCGTCTCTCTCACGGCG	259

Db	260	TCCATCCAGCTTGAGAGCTTAAGGCAATTAAACCCCTCCAGAAACAGCCAGTGGATGTAATGGCACA	319
Oy	240	GTGATCGGTGGACAG	253

Db 320 GTGATCGTGACAG 333

[illegible]

ACCESSION BF581041 CI:11654753
VERSION BF581041.1
KEYWORDS EST.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.

AUTHORS: NIH-MGC <http://mgc.ncl.nih.gov/>
TITLE: National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL: Unpublished (1999)
COMMENT: c:Health;nb:45;Structure;nb;D

Email: cgapds-remail.nin.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Assayed by: Robert M. C. E. Corporation (RLNC)

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

```

FEATURES
source      1.661
            location/qualifiers
            High quality sequence stop: 645.
            FILE: ZENR002.10M: 0 COLUMN: 15

```

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4219698"

```

```
lab_host="DH10B (T1 phage-resistant)"
```


Db	1428	AATGCTCTTGTTGATGCTTTCCACACACTCTCTCCAGAAATGCGGCAATGCTCCTCAAGC	1487
QY	180	TCGATCCAGCTTGAGAGCTAAGGGATTAACTCCAGAAACAGCCAGCTGATGTAATGGCACA	239
Db	1488	TCGATCCAGCTTGAGAGAGGAGGAGTATATCTCCAGAAATAAACCAATGATGTAATGGCTCA	1547
QY	240	GTGATCGTGAGCAG	253
Db	1548	GTGATCGTGAGCAG	1561
RESULT	10		
LOCUS	BF580957		
DEFINITION	BF580957 747 bp mRNA EST 12-DEC-2000		
ACCESSION	602100659F1 NCI_CGAP_Co24 Mus musculus CDNA clone IMAGE:4223990 5'		
VERSION	BF580957		
KEYWORDS	BF580957.1 GI:11654669		
SOURCE	EST.		
ORGANISM	house mouse.		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
JOURNAL	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
COMMENT	1 (bases 1 to 747) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov plate: LLAM9813 row: e column: 15 High quality sequence stop: 615. Location/Qualifiers 1. 747 /organism="Mus musculus" /strain="PVB/N" /db_xref="taxon:10090" /clone_image="IMAGE:4223990" /clone_id="NCI_CGAP_Co24" /lab_host="DH10B (T1 phage-resistant)" /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 Kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."		
BASE COUNT	203 a 198 c 187 g 159 t		
ORIGIN			
Query Match	55.0%; Score 139.2; DB 11; Length 747;		
Best Local Similarity	78.5%; Pred. No. 2.2e-31;		
Matches 179; Conservative	0; Mismatches 48; Indels 1; Gaps 1;		
QY	1	AACAAAGTGGGCCATCATCCACAGTCGCTTTGGGGCCCTGAGTCAGAAACTAG	60
Db	278	AGCAGAGGGGGGCGATCATCTAAGAGTGGCCCTGGAGCCGCTGCTTAAGAGACTTG	337
QY	61	AGGAGCTCTCCAAAATGACAGAGAGGTTTACAGACATATCTTCAGTCAAGTTACAGACA	120
Db	338	AGCAGCTCTCCAAAATGACAGAGAGGCGTCGACACATACCTTCGATCAGTTTCAGACA	397
QY	121	ATGGCCCTCATGTAGCTTTTGGGGCCCTTTCATCAGACAAATGAGAGCTGTCTCAGCGCT	180
Db	398	ATGCTCTTGTGTATGCTTTCGACGACACTCTCTCAGGAATGCGGCGATGCTCAGACT	457
QY	181	CCATCCAGCTTGAG--AGTAAAGGATTAACCCGCCGAAGAGCAGTGG	227
Db	458	CAATCCAGCTTGAGCAGCAGGGGAGTATATCTCCAGAAATTAACCAATGG	505

RESULT	11	831 bp	mRNA	EST	21-FEB-2001
LOCUS	BG286204				
DEFINITION	602383026p1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4500439 5',				
ACCESSION	BG286204				
VERSION	BG286204.1				
KEYWORDS	EST.				
SOURCE	human;				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 831)				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10366 row: d column: 08 High quality sequence stop: 685.				
FEATURES	Location/Qualifiers				
SOURCE	1..831				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:4500439"				
	/clone_lib="NIH_MGC_93"				
	/tissue_type="translational cell papilloma, cell line"				
	/lab_host="DH10B (phage-resistant)"				
	/note="Organ: Bladder; Vector: PCMV-SPORE6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."				
BASE COUNT	246 a 174 c 192 g 219 t				
ORIGIN					
	7 Query_Match 51.7%; Score 130.8; DB 11; Length 831;				
	Best local similarity 71.8%; Pred. No.7.5e-29;				
	Matches 171; Conservative 0; Mismatches 67; Indels 0; Gaps 0;				
CY	1 AACAAAGTGGGCCATCATCCACACAGTCGCTTGGGGCCCTCGCAGCTCAAGAAGTAG 60				
DB	367 AACAAAGTGGGCCATCATTCATTTATTCCTTTGGGAAGACTGCTGATGAAGCAGTAA 406				
CY	31 AGGAGCTGTCCAAATGACAGAGAGGTTTACAGACATATCTTCAGATCAAGTTTCAGACA 120				
DB	407 TAGAGATGAGCAAGATTAACAGAGAGGAAGCATTTTATTTGATTCAGATGAAGCTCAGAAC 466				
CY	121 ATGGCTCATATGATGCTTTTGGGGCCCTTTCATCATCAGAAATGGAGCTGCTCTCAGCGCT 180				
DB	457 ATGGCTCATATGATGCTTTTGGGGCCCTTTCATCATCAGAAATGGAGCTGCTCTCAGCGCT 526				
CY	181 CCATCCACCTTGAGAGTAAGGATTAACCCCTCCACAAACAGCAGGATGATGAATGGCAGC 238				
DB	527 CCCTTCACCTGGAAGGATTAAGGATTAACACACTGATATGTAATGCTCGATGAACGACAC 584				
RESULT	12	704 bp	mRNA	EST	12-JUN-2001
LOCUS	BG965023				
DEFINITION	602831666p1 NCI_CGAP_CO24 Mus musculus cDNA clone IMAGE:4986281 5',				
ACCESSION	BG965023				

VERSION	BC965023.1	GI:14352660
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ .	
AUTHORS	1 (bases 1 to 704)	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMN at: http://image.lnl.gov plate: L1AM10995 row: g column: 18 High quality sequence stop: 663.	
FEATURES	Location/Qualifiers	
source	1..704	
	/organism="Mus musculus"	
	/strain="FVB/N"	
	/db_xref="taxon:10090"	
	/clone="IMAGE:4986281"	
	/clone_lib="NCI CGAP Co24"	
	/lab_host="DH10B (TI phage-resistant)"	
	/note="Organ: colon; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."	
BASE COUNT	179 a	174 c 189 g 162 t
ORIGIN		
Query Match	50.0%	Score 136.4; DB 11; Length 704;
Best Local Similarity	76.4%	Pred. No. 1.5e-27;
Matches 194; Conservative	0; Mismatches 56;	Indels 4; Gaps 3
QY	1 AACAAAGTGGTCCATCATCCACACAGTCGCTTGGGGCCCTCTG-CAGCTCAAGAAGCTA 59	
DB	391 ACCAGAGCGGGGCATCATTCATACAGTGGCCCTGGGACCGCTGCCCTTAAAGAGCTT 450	
QY	60 GAGGAGCTGTCCAAATGACAGAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAAGAC 119	
DB	451 GAGCAGCTGTCCAAATGACAGAGAGGCTTCGACAGACATCTTCGGATGAGGTTCAAGAC 510	
QY	120 AATGCGCTATTCATTCCTTTGGGGCCCTTTCATTCAGGAATGAGAGCTGTCTCTCAGCGC 179	
DB	511 AATGCTCTTGTTCATCTTTCGACAGCTCTCTCTCGAAGAAATGCGGATCGTTCAGCAC 570	
QY	180 TCCATCCAGCTTGAGAGATA-AGGAGTTAACCCTCCAGACAGC--CAGTGAGTAATGGC 236	
DB	571 TCCATCCAGCTTGAGAGAGGAGGAGTATATCTCCAGAAATACCAACATGATGATGATGCG 630	
QY	237 ACAGTATCTGTGA 250	
DB	631 TCCGTGATCTGTGA 644	
RESULT 13		
LOCUS	BF355533	550 bp mRNA
DEFINITION	RC3-H10865-030800-021-b01 H10865 Homo sapiens	EST
ACCESSION	BF355533	22-NOV-2000
VERSION	BF355533.1	GI:11314594
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	

REFERENCE	TITLE	JOURNAL	MEDLINE	COMMENT
(bases 1 to 550)				
DIAS NETO, E., GARCIA CORTEA, R., VERJOVSKI-ALMEIDA, S., BRIONES, M.R., NEGAL, M.A., DA SILVA, W. JR., ZAGO, M.A., BORDIN, S., COSTA, F.F., GOLDMAN, G.H., CARVALHO, A.F., MATSUMURA, A., BALA, G.S., STIMPSON, D.H., BRUNSTEIN, A.H., DEOLIVEIRA, P.S., BUCHER, P., JONGENEEL, C.V., O'HARE, M.J., SOARES, F., BRENTANI, R.R., REIS, L.F., DE SOUZA, S.J. and STIMPSON, A.J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A.	97 (7), 3491-3496 (2000)	
20202663				
Contact: Stimpson A.J.G.				
Ludwig Institute for Cancer Research				
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil				
Tel: +55-11-2704922				
Fax: +55-11-2707001				
Email: asimpson@ludwig.org.br				
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:				
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2=RC3-HT0865)				
Seq primer: puc 18 forward				
High quality sequence start: 71				
High quality sequence stop: 548.				
Location/Qualifiers				
1..550				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/clone_11b="HT0865"				
/dev_stage="Adult"				
/note="Organ: head/neck; Vector: puc18; Site: 1: Smal; Site: 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) Profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."				
low stringency conditions."				
BASE COUNT	148 a	111 c	127 g	164 t
ORIGIN				
Query Match	41.1%	Score 104;	DB 11;	Length 550;
Best Local Similarity	74.4%	Pred. No. 7,4e-21;		
Matches 131; Conservative	0;	Mismatches 45;	Indels 0;	Gaps 0
Db	78	ACAGAGGTTTACAGACATATGCTTCAGATTCACAGATTCAGAACAAATGGCCCTCATTTGATGCT	137	
QY				
Db	549	ACAGAGGAGAGCATATTTATTTATGTTTCAGATGAAGCTCAGAGCAATGGCCCTCATTTGATGCT	490	
QY	138	TTTGGGGCCCTTCATCAGAGAAATGAGCGTCTCTCAGCGCTCCATCCACGCTTGAGAGT	197	
Db	489	TTTGGGGCTTCTTACACAGAGAAATGATGATCTCTCCAGAAAGTCCCTTCAGCTCGAAGT	430	
QY	138	AAGGAGTTAACCTCCAGAACAGCCAGTCGATGATGCGACAGTCACTGCGAGACG	253	
Db	429	AAGGAGTTAACCTGATATGTAATGCTGATGATGACAGACACTCTCATTAATGATAG	374	
RESULT 14				
LOCUS	BF581342	906 bp	MRNA	EST
DEFINITION	60210072661 NC_1_GAP_CO24 Mus musculus cDNA clone IMAGE:422935 5'			12-DEC-2000
ACCESSION	BF581342			
VERSION	BF581342.1	GI:11655054		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	1 (bases 1 to 906)			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			

This Page Blank (uspto)

STATE: NO. 6057492th Carolina
COUNTRY: US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-0900
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 9723 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 10..7419
US-08-083-590A-21

Query Match 11.3%; Score 28.6; DB 1; Length 9723;
Best Local Similarity 49.0%; Pred. No. 5;
Matches 76; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Dy 11 TGCATATCCACACAGTCGCTTTGGGGCCTTTCGACCTCAACAAGTAGAGACTGTG 70
||| | | | | | | | | | | | | | |
Db 2221 TGAGACCCTCCAGCATGATTCCATGATGACAGGAGATTGCTCGATTCCTCACCTGTG 2162

Dy 71 CAATAATGACAGAGATTATACAGACATATGCTTGATGATCAGTAAGTTAGAACAATGCCCTCAT 130
||| | | | | | | | | | | | | | |
Db 2161 AGTACACCTGGGGTATGGGCTCCCTCGGGGCAATATACAGCGGAACCATTCACACCGT 2102

Dy 131 TGATGCTTTGGGGCCCTTTTCATCAGGAATGAG 165
||| | | | | | | | | | | | | | |
Db 2101 TGATACATGTTGCACACCTTGGCAGCAGGAGATTGAG 2067

RESULT 4
US-08-532-384-21/c
Sequence 21, Application US/08532384
Patent No. 6083904

GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,384
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,590
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-0900
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:


```

? LENGTH: 9723 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: unknown
? MOLECULE TYPE: CDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 10..7419
?
US-08-532-384-21

```

Query Match	11.3%	Score	28.6	DB	3	Length	9723
Best Local Similarity	49.0%	Pred.	No. 5				
Matches	76	Conservative	0	Mismatches	79	Indels	0
						Gaps	0

Qy	11	TGCCATATCCACACAGTCCCTTTGGGGCCCTCCGACGCTCAAGAAACATGAGAGACCTTC	70
Db	2221	TGAGCTCCACAGTACAGTTTCCATGGATCAGAGGATGGCTCAGGCAATTCGTCACACTGTG	2162
Qy	71	CAAAATCAGAGGAGTTTACAGCATATGCTTCAGTCAAGTTCAGAAACATGGCCCTAT	130
Db	2161	AGTAGCAGCTGGGGTGATGGGGTCCCTCGGGCATTTACAGGGGAACCATTCACACCTT	2102
Qy	131	TGATGCTTTTGGGGCCCTTTCATCAGAGAAATGAG	165
Db	2101	TGATACATGTTGACCCCTTCGCGACAGGAGATGGAG	2067

RESULT 5
US-08-594-031-80

```

: GENERAL INFORMATION:
: APPLICANT: THOMPSON, Timothy C.
: TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
:
: NUMBER OF SEQUENCES: 175
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BAKER & BOTS, L.L.P.
: STREET: 1299 Pennsylvania Avenue, N.W.
: CITY: Washington

```

```

1 ZIP: 20004-2400
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Diskette
6
7 COMPUTER: IBM Compatible
8
9 OPERATING SYSTEM: DOS
10
11 SOFTWARE: FastSeq Version 1.5
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: US/08,594,031
16
17 FILING DATE: 30-JAN-1996

```

INFORMATION FOR SEQ ID NO: 80:

```
?
? SEQUENCE CHARACTERISTICS
? LENGTH: 371 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHEetical: NO
? ANTI-SENSE: NO
? FRAGMENT TYPE:
```

ORIGINAL SOURCE
US-08-594-031-80

```

QueryMatch      11.1%;  Score 28;  DB 1;  length 371;
Best Local Similarity 55.0%;
Matches 55;  Conservative 0;  Mismatches 45;  Indels 0;  Gaps 0;

```

0Y GATCGTTTGGGCGCCCTTCATCAGGAATAAGSACTGTCT 171
 1D TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
 1A3 AGAATGTCTTCTTGTCACATCCAGAACTGGAAATGTTTT 174

US-08-952-284-3/c

GENERAL INFORMATION:
 APPLICANT: Luna, Elizabeth J.
 APPLICANT: Pestonjamsap, Kersl N.
 APPLICANT: Pope, Robert K.
 APPLICANT: Wilkhuete, Julia D.
 TITLE OF INVENTION: ACTIN-BINDING POLYPEPTIDES
 TITLE OF INVENTION: AND NUCLEIC ACIDS ENCODING THE SAME
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston

```

ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
*COMPUTER: IBM Compatible
*OPERATING SYSTEM: Windows95
*SOFTWARE: FASTSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/962,284
Filing date: 11/08/96-1007

```

;: INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

SEQUENCE CHARACTERISTICS:
LENGTH: 6463 base pairs
%TYPE: nucleic acid
STRANDEDNESS: double
%POPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

Query Match	11.0%	Score 27.8	DB 2	Length 6463
Best Local Similarity	55.8%	Pred. No. 7.7		
Matches 53	Conservative 0	Mismatches 42	Indels 0	Gaps 0

cy 138 AATGAGCTGTCCTCAGCGCTCATCCAGCTTGAGAGTAAGGATTAACCTCCAGAA 217
Db 3921 ATATGCCCACTGCGCGCCCTGCTTCCTTGGAATGAGGGATTCATTCCCAAG 38622

QY 218 CAGCAGTGAATGACAGAGTATCGTGACA 252
| | | | | | | | | | | | | | | | | |
Db 3861 GATCAGCGGTTAATGTACAAATTCGTAGTCA 3827

RESULT 7

US-08-331-081B-4/C
Sequence 4, Application US/08331081B
Patent No. 5998697
GENERAL INFORMATION:
APPLICANT: Devlin, Robert H.
TITLE OF INVENTION: Transgenic Fish and Vectors Therefor
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: McFadden, Fincham
STREET: 225 Metcalfe Street, Suite 606
CITY: Ottawa
STATE: Ontario
COUNTRY: Canada
ZIP: K2P 1P9
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM PC or Compatibles
OPERATING SYSTEM: PC DOS 5.0
SOFTWARE: Wordperfect 5.1 (Patseq.txt)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331,081B
FILING DATE: October 20, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Canada 2,126,138
FILING DATE: June 17, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Fincham, H. Ian
REGISTRATION NUMBER: 26,375
REFERENCE/DOCKET NUMBER: 5478-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 234-1907
TELEFAX: (613) 234-5233
INFORMATION FOR SRO ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3872 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-331-081B-4

Query Match 10.9%; Score 27.6; DB 2; Length 3872;

Best local Similarity 48.1%; Pred. No. 6.9;

Matches 78; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 46 CAGCTCAGAGTCTGAGAGCTGTCACAAATGACAGAGCTTTACAGACATATGCTTCAG 105
| | | | | | | | | | | | | | | | | |
Db 2213 CACATTCAGGCTGCAATGAGTTGTCACAAATTCAGTACCTGTATTTATTTCTTGG 2154
QY 106 ATCAAGTTCAAGAACATGAGCTTCATGATGCTTTGGGGCCCTTCATCAGGAATGAG 165
| | | | | | | | | | | | | | | | | |
Db 2153 TGAATAATGTAGAAAATGTACGTATATCTGTGTAGCATTAAGATGAATGAATGAAT 2094
QY 166 CTGCTCTCAGCGCTGCATCCAGCTTGAGAGTAAAGGATTAA 207
| | | | | | | | | | | | | | | | | |
Db 2093 CAATCACTCAATGTACGTATCATGCAAAAACACAGATTATTA 2052

RESULT 8

US-08-486-013-65
Sequence 65, Application US/08486013
Patent No. 5731149
GENERAL INFORMATION:
APPLICANT: Seistled, Michael E.
ADDRESSEE: Ouellette, Andre J.

TITLE OF INVENTION: Antibiotic Cryptidin Peptides and Methods
TITLE OF INVENTION: of Their Use
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,013
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/342,268
FILING DATE: 18-NOV-1994
APPLICATION NUMBER: US 07/930,649
FILING DATE: 14-AUG-1992
PREOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,020
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UC 1206
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SRO ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-486-013-65

Query Match 10.7%; Score 27; DB 1; Length 445;

Best local Similarity 49.6%; Pred. No. 3.7;

Matches 69; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 93 ACATATGCTTGAATGATGATGACAAACATGCGCTCATGCTTTGGGGCCCTTTCA 152
| | | | | | | | | | | | | | | | | |
Db 18 ACCAATCCTCCAGGTGACCTTCACACCATGAAGACTCTGCTCCTGCGCCCTTGC 77
QY 153 TCAGGAATGGAGCTGTCTCCAGCGCTCCATCCAGCTTGAGAGTAAAGGATTAACTTC 212
| | | | | | | | | | | | | | | | | |
Db 78 TGTGGCATTTCCAGTCCAGGCTCATCCATTCAGAGGCGAAGAGACTAAACTG 137
QY 213 CAGAACGCCAGTGATGA 231
| | | | | | | | | | | | | | | | | |
Db 138 AGGAGACGCCAGCATGA 156

RESULT 9

US-08-486-013-67
Sequence 67, Application US/08486013
Patent No. 5731149
GENERAL INFORMATION:
APPLICANT: Seistled, Michael E.
ADDRESSEE: Ouellette, Andre J.
TITLE OF INVENTION: Antibiotic Cryptidin Peptides and Methods
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,013
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/342,268
FILING DATE: 18-NOV-1994
APPLICATION NUMBER: US 07/930,649
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,020
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UC 1206
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-486-013-67

Query Match 10.7%; Score 27; DB 1; Length 445;
Best Local Similarity 49.6%; Pred. No. 3.7;
Matches 69; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 93 ACATATGCTTCAGATCAAGTTCAGAACATGCGCTTCATTCGTTTGGGCGCCCTTTCA 152
DB 18 ACCAATCCCTCCAGTGACTTCAGCCATGAAGACTCTTCCTCCCTGCGCCCTTGTCC 77
QY 153 TCAGGAATGAGAGCTGTCTCTCAGCGCTCCATCCAGCTTGAGAGTAAGGATTAACCCTC 212
DB 78 TCGTGGCATTCAGATCCAGCGCTGATCCCATTCAGAGGACAGAGAGACTAAACTG 137
QY 213 CAGAACAGCCAGTGATGA 231
DB 138 AGGACAGCCAGCAGATGA 156

RESULT 10
US-08-482-279-65
Sequence 65, Application US/08482279
Patent No. 5840498
GENERAL INFORMATION:
APPLICANT: Selsted, Michael E.
APPLICANT: Ouellette, Andre J.
TITLE OF INVENTION: Antibiotic Cryptidin Peptides and Methods
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,279
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/342,268
FILING DATE: 18-NOV-1994
APPLICATION NUMBER: US 07/930,649
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,020
FILING DATE: 26-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UC 1206
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-482-279-65

Query Match 10.7%; Score 27; DB 2; Length 445;
Best Local Similarity 49.6%; Pred. No. 3.7;
Matches 69; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 93 ACATATGCTTCAGATCAAGTTCAGAACATGCGCTTCATTCGTTTGGGCGCCCTTTCA 152
DB 18 ACCAATCCCTCCAGTGACTTCAGCCATGAAGACTCTTCCTCCCTGCGCCCTTGTCC 77
QY 153 TCAGGAATGAGAGCTGTCTCTCAGCGCTCCATCCAGCTTGAGAGTAAGGATTAACCCTC 212
DB 8 TCGTGGCATTCAGATCCAGCGCTGATCCCATTCAGAGGACAGAGAGACTAAACTG 137
QY 213 CAGAACAGCCAGTGATGA 231
DB 138 AGGACAGCCAGCAGATGA 156

RESULT 11
US-08-482-279-67
Sequence 67, Application US/08482279
Patent No. 5840498
GENERAL INFORMATION:
APPLICANT: Selsted, Michael E.
APPLICANT: Ouellette, Andre J.
TITLE OF INVENTION: Antibiotic Cryptidin Peptides and Methods
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,279
FILING DATE:

INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-342-268-67

Query Match 10.7%; Score 27; DB 2; Length 445;
Best Local Similarity 49.6%; Pred. No. 3.7;
Matches 69; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 93 ACATATGCTTCAGATCAAGTTCAGAACATGCGCTCATGATGCTTTGGGCGCTTTCA 152
DB 18 ACCAATCCCTCCAGGAGTCCAGCTCCAGCATGAAAGACTCTTCCCTCTGCGCTTGTC 77
QY 153 TCAGAAATGAGAGCTGTCTCTCAGCGCTCCATCCAGCTTGAGAGTAAGGATTAACCTC 212
DB 78 TCGTGCATTCAGATCCAGCTGATCCATTCAAGAGCAGAGAGACTAAACTG 137
QY 213 CAGACAGCCAGTGTGA 231
DB 138 AGAGCAGCCAGCAGATGA 156

RESULT 14
US-09-015-968-65
Sequence 65, Application US/09015968
Patent No. 6057425

GENERAL INFORMATION:
APPLICANT: Seistled, Michael E.
APPLICANT: Ouellette, Andre J.
TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
NUMBER OF SEQUENCES: 70

CORRESPONDENCE ADDRESS:
ADDRESS: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/015,968
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/482,279
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/342,268
FILING DATE: 18-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,649
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,020
FILING DATE: 26-MAY-1992

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UC 3003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:

LENGTH: 445 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-015-968-65

Query Match 10.7%; Score 27; DB 3; Length 445;
Best Local Similarity 49.6%; Pred. No. 3.7;
Matches 69; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 93 ACATATGCTTCAGATCAAGTTCAGAACATGCGCTCATGATGCTTTGGGCGCTTTCA 152
DB 18 ACCAATCCCTCCAGGAGTCCAGCTCCAGCATGAAAGACTCTTCCCTCTGCGCTTGTC 77
QY 153 TCAGAAATGAGAGCTGTCTCTCAGCGCTCCATCCAGCTTGAGAGTAAGGATTAACCTC 212
DB 78 TCGTGCATTCAGATCCAGCTGATCCATTCAAGAGCAGAGAGACTAAACTG 137
QY 213 CAGACAGCCAGTGTGA 231
DB 138 AGAGCAGCCAGCAGATGA 156

RESULT 15
US-09-015-968-67
Sequence 67, Application US/09015968
Patent No. 6057425

GENERAL INFORMATION:
APPLICANT: Seistled, Michael E.
APPLICANT: Ouellette, Andre J.
TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
NUMBER OF SEQUENCES: 70

CORRESPONDENCE ADDRESS:
ADDRESS: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/015,968
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/482,279
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/342,268
FILING DATE: 18-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,649
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,020
FILING DATE: 26-MAY-1992

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UC 3003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 base pairs
TYPE: nucleic acid

XX	PI	Nakanishi A, Morita S;
DR	WPI: 2001-355935/37.	
XX	P-PSDB; AAB73716.	
PT	New antisense nucleotide, useful for treatment and prevention of	
PT	bronchial asthma and chronic obstructive pulmonary disease	
XX		
PS	Claim 3; Page 82-84; 104pp; Japanese.	
XX		
CC	The invention relates to an antisense nucleotide targetted to the mouse	
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,	
CC	the CLCA1 gene (coding sequence shown in AAH46102). The invention also	
CC	relates to an antibody specific for the Gob-5 protein, medical and	
CC	diagnostic compositions containing the antisense nucleotide or the	
CC	antibody, and methods and kits for screening for compounds which inhibit	
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.	
CC	The antisense oligonucleotides and antibody are therefore useful for the	
CC	treatment and prevention of bronchial asthma and chronic obstructive	
CC	pulmonary disease. The present sequence represents the human CLCA1	
CC	gene coding sequence.	
SQ	Sequence 2742 BP: 833 A; 616 C; 623 G; 670 T; 0 other:	
Query Match	100.0%; Score 253; DB 22; Length 2742;	
Best Local Similarity	100.0%; Pred. No. 8,9e-74;	
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 AACAAAGTGTGCCATCATCCACAGTGTGCTTTGGGCGCTCTCGAGCTCAAGAACTAG 60	
Db	1277 aacaagaatgtgtccatcatccacacagtcgtcttgggccccttcgaagtcagaactag 1336	
OY	61 AGGACCTGTCCAAATGACAGGAGGTTTACACAAATGCTTCAGTWCAGTTCGAACA 120	
Db	1337 aggaagctgcacaatgaaagaggtttacaacaataagctcagaatcagttcagaaca 1396	
OY	121 ATGGGCTCATTTGATCTTTTGGGGCCCTTTTCATCAGAGAAATGAGACCTGTCTCAACGCT 180	
Db	1397 atggcctcatgtatgcttttggggcccttcacagaaatgagctgtctcagcgct 1456	
OY	181 CCATCCAGCTTAGAGTAAGGATTAAACCTCCAGAACGCCAGTGGATGGAATGGCAGAG 240	
Db	1457 ccataccagcttgagagtaaggattaacccctccagaacagacagtgagatgaaatggcagag 1516	
OY	241 TGATCGTGGACAG 253	
Db	1517 tgatcgtggacag 1529	
RESULT	2	
AAF81927	AAF81927 standard; cDNA: 2745 BP.	
XX	AAF81927;	
AC		
XX	13-JUN-2001 (first entry)	
DE		
XX	Human ICACC-1 nucleotide sequence.	
XX		
KW	ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;	
KW	interleukin 9 induced calcium activated chloride channel; IL-9;	
KW	calcium activated chloride channel; anti-allergic; anti-asthmatic;	
KW	anti-inflammatory; immunomodulatory; cystic fibrosis;	
XX	inflammatory bowel disease; autoimmune disease; ss.	
OS	Homo sapiens.	
PH	Key	Location/Qualifiers
FT	CDS	1..2745
FT		/tag= a
FT		/product= "ICACC-1"

```

FT      /note= "IL-9 induced calcium activated chloride channel"
PN      WO9944620-A1.
XX      10-SEP-1999.
PD      03-MAR-1999; 99WO-US04703.
XX      03-MAR-1999; 99US-0076815.
XX      03-MAR-1998; 98US-0076815.
PR      (MAGA-) MAGAININ PHARM INC.
PA      HoJroyd KJ, Levitt RC, Maloy WT, Louahed J, McLane M;
PI      Nicolaidis NC, Zhou Y, Dong Q;
XX      WPI, 1999-550979/46.
DR      P-PSDB; AAB74824.
XX      New nucleic acid encoding calcium activated chloride channel, used to
PT      identify, e.g. specific modulators for treating atopic allergy -
PS      Claim 1; Fig 4B; 75pp; English.
CC      The present sequence encodes the human interleukin 9 (IL-9) induced
CC      calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
CC      have anti-allergic, anti-asthmatic, anti-inflammatory and
CC      immunomodulatory activities. Compounds (A) that downregulate ICACC are
CC      used to alleviate asthma (or more generally atopic allergy), while those
CC      (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
CC      inflammatory bowel disease (IBD) (or other autoimmune diseases).
CC      Measuring levels of ICACC-1 can be used in the diagnosis of asthma
CC      (increased levels) or IBD (reduced levels), also for monitoring
CC      treatment of these conditions. The ICACC proteins can be used:
CC      (i) to raise specific antibodies (Ab), useful: (a) as immunoassay
CC      reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
CC      (A) to disrupt bonding between ICACC and its ligands; and (iii) to
CC      identify modulators and binding partners. ICACC polynucleotides can be
CC      used to generate transgenic animals or recombinant cells, used to screen
CC      for antagonists, also as a source of therapeutic antisense agents or
CC      diagnostic probes (for quantifying mRNA expression, e.g. for
CC      identification of modulators).
SQ      Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other:
QY      Query Match          100.0%; Score 253; DB 20; Length 2745;
QY      Best Local Similarity 100.0%; Pred. No. 8.9e-74;
QY      Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0
DB      1 AACMAAGTGTGTCATCATCACACAGTGGCTTGGGCGCCCTCTCAGCTCAAGAACTAG 60
DB      1277 aacaaagtgtgcatcaccacacagtcgcttggggcccttgagctcaagaactag 1336
QY      61 AGAGAGTGTCCAAATGACAGAGGAGTTTCAGACATATGCTTCAGATTCAGATTACAGAA 120
DB      1337 aggaagctgtccaaaagacaagagggtttcaagacacatagtctcaagtcgaagtcgaaca 1396
QY      121 ATGGCCCTCATGTATGCTTTTGGGGCCCTTTCATCAGAGAAATGAGAGCTGTCTCACCGCT 180
DB      1397 atggcctcatgtatgcttttggggcccttccatcaggaatgtagctgtctccagcgct 1456
QY      181 CCATCCAGCTTGAGAGTTAAGGATTAACTCCAGAACAGCCAGTGTGATGATGAGACAG 240
DB      1457 ccatacagcttgagagtaaggattaacctccaagaacagccagtgatgaatcagcag 1516
QY      241 TGATCGTGCAGCAG 253
DB      1517 tgatcgtgcagcag 1529

```


AAH46124;	11-SEP-2001	(first entry)
Human CLCA1 cDNA, SEQ ID NO:26.		
Human CLCA1: goblet cell; mouse Gob-5 orthologue; drug screening; expression inhibition; antisense therapy; gene therapy; chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic; ss.		
Homo sapiens.		
Key	Location/Qualifiers	
CDS	23..2767	
FT	/*tag= A	
FT	/product= "Human CLCA1"	
FT	/trans_except= (pos:476..478, aa:Lys)	
MO200138530-A1.		
31-MAY-2001.		
22-NOV-2000; 2000WO-JP08232.		
24-NOV-1999; 99JP-0333479.		
27-APR-2000; 2000JP-0127589.		
(TAKE) TAKEDA CHEM IND LTD.		
Nakanishi A, Morita S;		
WPI: 2001-355935/37.		
P-PSDB: AAB73716.		
New antisense nucleotide, useful for treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease		
Example 5; Page 92-94; 104pp; Japanese.		
The invention relates to an antisense nucleotide targeted to the mouse Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart, the CLCA1 gene (coding sequence shown in AAH46102). The invention also relates to an antibody specific for the Gob-5 protein, medical and diagnostic compositions containing the antisense nucleotide or the antibody, and methods and kits for screening for compounds which inhibit the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells. The antisense oligonucleotides and antibody are therefore useful for the treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease. The present sequence represents human CLCA1 cDNA.		
Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other:		
Query Match	100.0%;	Score 253; DB 22; Length 2825;
Best Local Similarity	100.0%;	Pred. No. 9e-74;
Matches 253; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
1 AACAAAGGAGGCAATCATCCACACAGTCGCTTTGGGGCCCTTCAGCTCAAGAATGAG 60		
1299 aacaaaggggagccatccacacagtcgcttggggcccttcgacgtcagaactag 1358		
61 AGGAGCTGTCCAAAATGACAGAGAGGTTTACAGACATATGCTTCAATCAAGTTTCAGAA 120		
1359 agggagctgtccaaaaatgacagagaggtttacagacataatgcttcagataatgacagaaca 1418		
121 ATGGCCCTCATTTGATGCTTTTGGGGCCCTTTTCATCATCAGAAATGAGACTGTCTTCAGCGCT 180		
1419 atggccctcatatgagcttttggggccctttcatcaggaatgagagctgtctctcagcgct 1478		
181 CCATTCAGACTTGAGAGTAAAGGATTAAACCTTCAGAACACAGCCAGTCGATGATGGCAGAG 240		

Do 1479 ccatccagcttgtaggaaggaattaacctccagaaacagcagtgatgtagtgcacag 1538
 Cy 241 TGATCGTGACAG 253
 |||||||
 Db 1539 tgatcgtgagacag 1551
 RESULT 4
 ID AAH34879
 AAH34879 standard; cDNA; 2854 BP.
 AC AAH34879;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
 XX
 KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 1; ss.
 XX
 OS Homo sapiens.
 XX
 PN NC0200122920-A2.
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000MO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 RI Ruben SM, Barash SC, Blrse CE, Rosen CA;
 XX
 DR WEI; 2001-235357/24.
 DR E-SSDB; AAG54474.
 XX
 PS Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 useful for preventing, diagnosing and/or treating colorectal cancers -
 Claim 1; Page 3462-3463; 9803pp; English.
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytosolic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N-7. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SO Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Oy	61	AGGACGTGTCGCAAAATGACAGAGGTTTTCACAGCATATCTTCAGATCAAGTTCACAAACA	120		
Db	1371	aggaaatctgccaaatagaaagaaagtcttcacagacataatgcttccagaatcaagttcaagaaca	1430		
Oy	121	ATGGCCTCATGTATGCTGTTTTGGGGCCCTTTCATCAGAGAAATGGAGCTGTCTCTCAGCGCT	180		
Db	1431	atggcctcatgtatgcttcttggggcccttcacagaatagagctgtctctccagcgct	1490		
Oy	181	CCATCAGGTTGAGTAGGGATTACCTCCAGAAACGCCAGTGAGTGAATGAGGCACAG	240		
Db	1491	ccatcagcttgaagataaggatlaaccctccagaacagccagtgagatgaatgagacag	1550		
Oy	241	TGATGCTGGACAG	253		
Db	1551	tgatgctggacag	1563		
RESULT 5					
ID	AAF81787	AAF81787 standard; cDNA; 2854 BP.			
XX	AAF81787;				
DT	12-JUN-2001	(first entry)			
DE	Human secreted protein gene 1 SEQ ID NO:11.				
XX					
KW	Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;				
KW	dermatological; immunosuppressive; antinflammatory; anti-HIV;				
KW	immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;				
KW	ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;				
KW	antialzheimer; antiparkinsonian; antimicrobial; vunerary; gene therapy;				
KW	immune disorder; hyperproliferative disorder; cardiovascular disease;				
KW	cancer; angiogenic disorder; neurological disorder; infectious disease;				
KW	wound healing; regeneration; chemotaxis; chromosome 1; ss.				
OS	Homo sapiens.				
PN	MO200112775-A2.				
XX					
PD	22-FEB-2001.				
XX					
PF	16-AUG-2000; 2000WO-US22325.				
XX					
PR	17-AUG-1999; 99US-0149182.				
XX					
PA	(HUMA-) HUMAN GENOME SCT INC.				
PI	Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;				
PI	Birsek CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;				
XX					
DR	WPI: 2001-147550/15.				
DR	P-PSDB; AAB74733.				
XX					
PT	Nucleic acids encoding 25 human secreted polypeptides, useful for				
PT	preventing, diagnosing and/or treating e.g. cancers, Parkinson's				
PT	disease and diabetic retinopathy -				
XX					
PS	Claim 1; Page 441; 485pp: English.				
XX					
CC	AAB74733 to AAF81817 encode the human secreted proteins given in AAB74733				
CC	to AAB74772. Human secreted proteins can have activities based on the				
CC	tissues and cells they are expressed in. Example of activities include:				
CC	immunomodulatory; antisclerotic; dermatological; immunosuppressive;				
CC	antinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;				
CC	vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;				
CC	anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and				
CC	vunerary. Human secreted proteins can be used in gene therapy and				
CC	vaccine. Human secreted protein nucleotide sequences (NAM1) and proteins				
CC	(PEP1) may be used in the prevention, diagnosis and treatment of diseases				
CC	associated with inappropriate polypeptide expression. For example, NAM1				
CC	and PEP1 may be used to treat disorders associated with decreased				

[illegible]

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
CC
XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;

Query Match 100.0%; Score 253; DB 22; Length 2867;
Best Local Similarity 100.0%; Pred. No. 9e-74;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAAAGTGGTCCATCATCAGACAGTGGCTTTGGGGCCCTCTGAGCTCAAGACTAG 60
DB 1315 aacaaagtgagcgcacacacacagctgcttggggccctctgagctcaagaactag 1374
QY 61 AGGAGCTGTCCAAATGACAGAGAGGTTTACAGACATATGCTTCAGATCAGTTCAGACAA 120
DB 1375 aggagctgtccaaatgacagagaggtttacagacatattctcagatcagaactag 1434
QY 121 ATGGCCTCATTTATGCTTTTGGGGCCCTTTTCATCAGAGAAATGAGCTGTCTCAGCGCT 180
DB 1435 atggcctcatgtatgcttttggggccctttcatcagaagaatgagctgtctcagcgct 1494
QY 181 CCATCAGCTTGAGAGTAAAGGATTAAACCTCCAGAAAGCCAGTGCATGATGAGCAGAG 240
DB 1495 ccatacagcttgagagtaaggaattaaacccccaagaacagtgagatgaatgagcag 1554
QY 241 TGATCGTGGACAG 253
DB 1555 tgatcggtgacag 1567

RESULT 7
AAH35019
ID AAH35019 standard; cDNA; 3109 BP.

AC AAH35019;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:2101.

KM Human; colon cancer; colon cancer antigen; diagnosis; detection;

KM colorectal carcinoma; ss.

OS Homo sapiens.

PN W0200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX

(HBM-) HUMAN GENOME SCI INC.

Ruben SM, Barash SC, Birse CE, Rosen CA;

WPI: 2001-235357/24.

P: SDB: AAG75614.

Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
useful for preventing, diagnosing and/or treating colorectal cancers -
Chim 1; Page 3587-3588; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
CC
XX Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match 100.0%; Score 253; DB 22; Length 3109;
Best Local Similarity 100.0%; Pred. No. 9.4e-74;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAAAGTGGTCCATCATCAGACAGTGGCTTTGGGGCCCTCTGAGCTCAAGACTAG 60
DB 1164 aacaaagtgagcgcacacacacagctgcttggggccctctgagctcaagaactag 1223
QY 61 AGGAGCTGTCCAAATGACAGAGAGGTTTACAGACATATGCTTCAGATCAGTTCAGACAA 120
DB 1224 aggagctgtccaaatgacagagaggtttacagacatattctcagatcagaactag 1283
QY 121 ATGGCCTCATTTATGCTTTTGGGGCCCTTTTCATCAGAGAAATGAGCTGTCTCAGCGCT 180
DB 1284 atggcctcatgtatgcttttggggccctttcatcagaagaatgagctgtctcagcgct 1343
QY 181 CCATCAGCTTGAGAGTAAAGGATTAAACCTCCAGAAAGCCAGTGCATGATGAGCAGAG 240
DB 1374 ccatacagcttgagagtaaggaattaaacccccaagaacagtgagatgaatgagcag 1403
QY 241 TGATCGTGGACAG 253
DB 1404 tgatcggtgacag 1416

RESULT 8
AA209840
ID AA209840 standard; cDNA; 3111 BP.

AC AA209840;

DT 26-NOV-1999 (first entry)

DE Human membrane spanning protein MSP-5 cDNA fragment 2.

KM Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;

KM neoplastic disorder; immunological disorder; reproductive disorder;

KM MSP-5; ds.

XX

XX OS Homo sapiens.
XX PN WO9946380-A2.
XX PD 16-SEP-1999.
XX PF 09-MAR-1999; 99WO-US05073.
XX PR 13-MAR-1998; 98US-0039064.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Tang YF, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;
XX PI Guegler KJ, Kaser MR, Baughn MR, Shah P;
XX DR WPI; 1999-551409/46.
XX DR P-PSDB; AAY33298.
XX PT New human membrane spanning proteins used to, e.g. prevent and treat
XX PS neoplastic disorders -
XX PS Example 1; Page 80-81; 81pp; English.
XX CC This invention describes novel human membrane spanning proteins (MSPs),
XX CC and the polynucleotides encoding them. The products of the invention are
XX CC used to diagnose, prevent and treat neoplastic, immunological and
XX CC reproductive disorders. This sequence encodes a human membrane spanning
XX CC protein MSP-5 fragment.
XX SQ Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match 100.0%; Score 253; DB 20; Length 3111;
Best Local Similarity 100.0%; Pred. No. 9,4e-74;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAAAGTGTGTCATTCACACAGTCGCTTTGGGCGCTTCAGCTCAAGACTAG 60
DB 1310 aacaaagtggtgcatcaccacacagtcgcttggggccctcagctcaagactag 1369
QY 61 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCATCAAGTTCAAGACA 120
DB 1370 aggaagctgtccaaatgacagaggtttacagacatattgcttcagatcaagttcgaaca 1429
QY 121 ATGGCCTCATTTGATGCTTTGGGCGCTTTCATCAGGAATGAGACTGCTCTCAGCGCT 180
DB 1430 atggcctcatgttgccttggggcccttcctcaggaatgagactgctctcagcgct 1489
QY 181 CCATCCAGCTTGAAGAGTAAGGATTAACCTCCAGAACAGCCAGTGATGAATGGCAGC 240
DB 1490 ccattccagcttgagagtaaggttaaacctccagaaacagccagtgatgaatggcacag 1549
QY 241 TGAATCTGGACAG 253
DB 1550 tgatctggacag 1562

RESULT 9
AAI29502
ID AAI29502 standard; cDNA; 3311 BP.
XX AAI29502;
XX 12-OCT-2001 (first entry)
XX C902P determined cDNA sequence.
XX Human; immunotherapy; diagnosis; colon cancer; colon tumour;
XX Immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX Homo sapiens.

PN WO200149716-A2.
XX 12-JUL-2001.
XX 29-DEC-2000; 2000MO-US35596.
XX 30-DEC-1999; 99US-0476296.
XX 10-JAN-2000; 2000US-0480321.
XX 15-FEB-2000; 2000US-0504629.
XX 06-MAR-2000; 2000US-0519444.
XX 19-MAY-2000; 2000US-0575251.
XX 23-JUN-2000; 2000US-0609448.
XX 28-AUG-2000; 2000US-0649811.
XX (CORI-) CORIXA CORP.
XX PI Xu J, Lodes MT, Secret H, Benson DR, Meagher MJ, Stolk JA;
XX PI King GE, Wang T, Jiang Y;
XX DR WPI; 2001-441847/47.
XX PT Colon tumor associated proteins and nucleic acids useful for the
XX PS prevention, diagnosis and treatment of colonic cancer -
XX PS Claim 2; Page 425-426; 472pp; English.
XX CC This present invention describes colon tumor associated proteins (I) and
XX CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
XX CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
XX CC (II) may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate colon tumor associated protein (TCAP)
XX CC expression, such as colonic cancer. For example, (I) and (II) may be
XX CC used to treat disorders associated with decreased expression by
XX CC rectifying mutations or deletions in a patient's genome that affect the
XX CC activity of TCAPs by expressing inactive proteins or to supplement the
XX CC patient's own production of them. Additionally, (II) may be used to
XX CC produce the TCAP proteins, by inserting the nucleic acids into a host
XX CC cell culturing the cell to express the protein. (II) and its
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC polymerase chain reaction (PCR) and hybridisation assays to detect and
XX CC quantitate the presence of similar nucleic acids in samples, and
XX CC also be used as antigens in the production of antibodies against TCAPs
XX CC and in assays to identify modulators of TCAP expression and activity.
XX CC Anti-(I) antibodies and antagonists may also be used to down regulate
XX CC TCAP expression and activity. The anti-(I) antibodies may also be used
XX CC as diagnostic agents for detecting the presence of TCAPs in samples
XX CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
XX CC and AAM24494 to AAM24523 represent nucleotide and amino acid sequences
XX CC given in the exemplification of the present invention.
XX SQ Sequence 3311 BP; 1028 A; 692 C; 742 G; 849 T; 0 other;

Query Match 100.0%; Score 253; DB 22; Length 3311;
Best Local Similarity 100.0%; Pred. No. 9,6e-74;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAAAGTGTGTCATTCACACAGTCGCTTTGGGCGCTTCAGCTCAAGACTAG 60
DB 1638 aacaaagtggtgcatcaccacacagtcgcttggggccctcagctcaagactag 1687
QY 61 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCATCAAGTTCAAGACA 120
DB 1698 aggaagctgtccaaatgacagaggtttacagacatattgcttcagatcaagttcgaaca 1747
QY 121 ATGGCCTCATTTGATGCTTTGGGCGCTTTCATCAGGAATGAGACTGCTCTCAGCGCT 180
DB 1748 atggcctcatgttgccttggggcccttcctcaggaatgagactgctctcagcgct 1807
QY 181 CCATCCAGCTTGAAGAGTAAGGATTAACCTCCAGAACAGCCAGTGATGAATGGCAGC 240
DB 1808 ccattccagcttgagagtaaggttaaacctccagaaacagccagtgatgaatggcacag 1867

QY 241 TGATCGTGACAG 253
DB 1868 tgatcgtgacag 1880

RESULT 10
AAH46101
ID AAH46101 standard; DNA; 2739 BP.
XX
AC AAH46101;
XX
DT 11-SEP-2001 (first entry)
XX
DE Mouse Gob-5 coding sequence, SEQ ID NO:3.
XX
KM Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
expression inhibition; antisense therapy; gene therapy; bronchial asthma;
chronic obstructive pulmonary disease; antiasthmatic; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 1..2739
FT /tag=a
FT /partial
FT /product="Mouse Gob-5"
FT /note="No stop codon given in the specification"
XX
PN WO200138530-A1.
XX
PD 31-MAY-2001.
XX
PE 22-NOV-2000; 2000MO-JP08232.
XX
PR 24-NOV-1999; 99JP-0333479.
PR 27-APR-2000; 2000JP-0127589.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Nakanishi A, Morita S;
XX
DR WPI: 2001-355935/37.
DR P-PSDB; AAB73715.
XX
PT New antisense nucleotide, useful for treatment and prevention of
PT bronchial asthma and chronic obstructive pulmonary disease -
XX
PS Claim 3; Page 80-82; 104pp; Japanese.
XX
CC The invention relates to an antisense nucleotide targeted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the Gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
CC pulmonary disease. The present sequence represents the mouse Gob-5
CC gene coding sequence.
XX
SQ Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other;

Query Match 69.0%; Score 174.6; DB 22; Length 2739;
Best Local Similarity 80.6%; Pred. No. 8,4e-48;
Matches 204; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 AACAAAGTGTGCATCATCACAGAGTTCGCTTGGGCCCTCTGCGAGTCAAGACTAG 60
DB 1280 agcagagcgggcattcatcattacgtgcccctgggacgcgtgcgtctaaagactg 1339

QY 131 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTTCAGACA 120
DB 1340 agcagctgtccaaatgacagagggcctgcagacattcttcgattcaggttcagaaca 1399
QY 1421 ATGGCCTCATTTGATGCTTTGGGGCCCTTCAATCAGAAATGAGCTGTCTCCAGCGCT 180
DB 1490 atggtctctgtatgtcttcgcagcactctcccgaaatgagtgatgcctcagcact 1459
QY 1491 CCATCCAGCTTGAAGTAGGATTAACCTCCAGAACGCCAGTGTGATGAATGCGACAG 240
DB 1450 ccattcagcttgagagcagcaggaatgaattatctccagataaccatgagatgctcag 1519
QY 241 TGATCGTGACAG 253
DB 1520 tgatcgtgacag 1532

RESULT 11
AAH46120
ID AAH46120 standard; cDNA; 2843 BP.
XX
AC AAH46120;
XX
DT 11-SEP-2001 (first entry)
XX
DE Mouse Gob-5 cDNA, SEQ ID NO:22.
XX
KM Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
expression inhibition; antisense therapy; gene therapy; bronchial asthma;
chronic obstructive pulmonary disease; antiasthmatic; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 15..2756
FT /tag=a
FT /product="Mouse Gob-5"
XX
PN WO200138530-A1.
XX
PD 31-MAY-2001.
XX
PE 22-NOV-2000; 2000MO-JP08232.
XX
PR 24-NOV-1999; 99JP-0333479.
PR 27-APR-2000; 2000JP-0127589.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Nakanishi A, Morita S;
XX
DR WPI: 2001-355935/37.
DR P-PSDB; AAB73715.
XX
PT New antisense nucleotide, useful for treatment and prevention of
PT bronchial asthma and chronic obstructive pulmonary disease -
XX
PS Example 1; Page 89-91; 104pp; Japanese.
XX
CC The invention relates to an antisense nucleotide targeted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the Gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
CC pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
XX
SQ Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;

	Query Match	69.0%	Score 174.6:	DB 22:	Length 2843;	
	Best Local Similarity	80.6%;	Pred. No. 8.6e-48;			
	Matches 204; Conservative	0;	Mismatches 49;	Indels	0;	Gaps
OY	1 AACAAAGTGTGCGCATATCATCCACACACTGCCCTTTGGGGCCCTCTGCAGCTCAAGAACTAG 60					
Db	1294 agcagaggggcccatacatccatcaccttgccccctgggaaccgcgtcccgctaagaagccttg 1353					
OY	61 AGGAGCTGTCCAAAATGCACAGAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAGAACAA 120					
Db	1354 agcagcgtcccaaaatgacagagaggcctgagaaataccttccttgcagttcagaaca 1413					
OY	121 ATGGCCTATTGATGATGCTTTTGGGGCCCTTTCATCAGGAATAAGAGCTGTCTCACGCGT 180					
Db	1414 atgtctctgttgttagctcttcgcagcaactctccctcagaagaatgsgcgtcgcagcact 1473					
OY	181 CCATCCACGCTTGGAGATGAGGAGTAACCTCCAGAAACACCACGATGATGAATGCGACAG 240					
Db	1474 ccattccaagtctggagagcagggagtaatactccagaataaccaatgatgatgacctcag 1533					
OY	241 TGATGCTGACAG 253					
Db	1534 tgatcgtctggacag 1546					
	RESULT 12					
ID	AAF81925					
XX	AAF81925 standard; cDNA; 2931 BP.					
AC	AAF81925;					
DX	13-JUN-2001 (first entry)					
DT						
XX						
DE	Murine ICAC-1 nucleotide sequence.					
XX						
KM	ICAC-1; ICAC-2; asthma; atopic allergy; asthma-related disorder;					
KW	interleukin 9 induced calcium activated chloride channel; IL-9;					
KW	calcium activated chloride channel; anti-allergic; anti-asthmatic;					
KW	anti-inflammatory; immunomodulatory; cystic fibrosis;					
OS	inflammatory bowel disease; autoimmune disease; ss.					
XX						
OS	Mus sp.					
XX						
FH	Key	Location/Qualifiers				
FT	CDS	8..2749				
FT	/tag= a					
FT	/product= "ICAC-1"					
FT	/note= "IL-9 induced calcium activated chloride channel"					
XX						
PN	W09944620-A1.					
XX						
PD	10-SEP-1999.					
XX						
PF	03-MAR-1999; 99WO-US04703.					
PR	03-MAR-1998; 98US-0076815.					
XX						
PA	(MACA-) MAGAININ PHARM INC.					
XX	Holroyd KJ, Levitt RC, Maloy WL, Lounshed J, McLane M;					
PI	Nicolaides NC, Zhou Y, Dong Q;					
XX						
DR	WPI: 1999-550979/46.					
XX	P-PSDB; AAB74822.					
PT	New nucleic acid encoding calcium activated chloride channel, used to					
XX	identify, e.g. specific modulators for treating atopic allergy -					
PS	Claim 2; Fig 2; 75pp; English.					
XX						
CC	The present sequence encodes the murine interleukin 9 (IL-9) induced					
CC	calcium activated chloride channel 1 (ICAC-1) protein. ICAC proteins					
CC	have anti-allergic, anti-asthmatic, anti-inflammatory and					

immunomodulatory activities. Compounds (A) that downregulate ICACC are used to alleviate asthma (or more generally atopic allergy), while those (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and inflammatory bowel disease (IBD) (or other autoimmune diseases).
 Measuring levels of ICACC-1 can be used in the diagnosis of asthma (increased levels) or IBD (reduced levels), also for monitoring treatment of these conditions. The ICACC proteins can be used:
 (1) to raise specific antibodies (Ab), useful: (a) as immunoassay reagents, and (b) as therapeutics (A); (ii) as fragments or mutants, as (A) to disrupt bonding between ICACC and its ligands; and (iii) to identify modulators and binding partners. ICACC polynucleotides can be used to generate transgenic animals or recombinant cells, used to screen for antagonists, also as a source of therapeutic antisense agents or diagnostic probes (for quantifying mRNA expression, e.g. for identification of modulators).
 Sequence 2931 BP; 867 A; 719 C; 684 G; 661 T; 0 other;
 Query Match 69.0%; Score 174.6; DB 20; Length 2931;
 Best Local Similarity 80.6%; Pred. No. 8.7e-48;
 Matches 204; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY	1	AACAAAGTGTGCCATCATCCACAGATGCGCTTTGGGGCCCTCTGCACTCAAGAACTAG	60
DB	1287	agcagagcgggccatccatccatagcagtgcccgcgccgctgcgttaagagcttg	1346
QY	51	AGGACCTGTCGCAAAATGACAGCAGCGCTTACAGACATATGCTTCAGATCAAGTTCCAGACA	120
DB	1317	agcagcgtgccaaatagcagagcgccgcgcagacatactcttcgatacgtgttcgaaca	1406
QY	121	ATGGCCTCATTTATGCTTTTGGGCGCCCTTTATCATCAGAAATGGACTGTCTCTCAGCGCT	180
DB	1407	atggtcttgctatgctcttcgcagcactctctcagaaatgctgcgtcgcagact	1466
QY	181	CCATCCAGCTTGAGAGTAAAGGATTAAACCCCTCCACAAACAGCAGGAGATGATGGCAGAG	240
DB	1457	caatcagctgtgagagcgacggaggaatctccagaataaccatcgatgatgctcag	1526
QY	241	TGATCGTGACAG	253
DB	1527	tgatcgtgacag	1539

RESULT: 13
 AAAG4335
 ID AAAG4335 standard; DNA; 2616 BP.
 AC AAAG4335;
 XX
 XX 20 DEC-2000 (first entry)
 XX
 XX
 DE Clone 251688 of a novel gene associated with colon disease.
 XX
 XX
 KW Colon disease; colon cancer gene; carbonic anhydrase; galectin;
 KW carcinoembryonic antigen; colorectal carcinoma tumour associated antigen;
 KW fatty-acid binding protein; glutathione peroxidase; quanylin;
 KW cyclokeratin; cadherin; intestinal mucin; colon cancer; Crohn's disease;
 KW metastatic colon cancer; atrophic gastritis; cholecystitis;
 KW irritable bowel syndrome; ulcerative colitis; ss.
 XX
 OS Homo sapiens;
 XX
 XX MOJ00050588-A2.
 XX
 PD 31-AUG-2000.
 XX
 PF 01-FEB-2000; 2000MO-DS02595.
 XX
 PR 22-FEB-1999; 99US-0255381.
 XX
 PA (INCYT-) INCYTE PHARM INC.
 XX

PI Walker MG, Volkmuth W, Klingler TM, Lal P;
XX WPI: 2000-558397/51.
DR
XX Novel gene associated with colon cancer and coexpressed with one or
PT more known colon cancer genes in a number of biological samples, for
PT use in gene therapy
XX
PS Claim 2: Page 30-31; 33pp; English.
XX
CC The present sequence represents a novel gene which is associated with
CC colon disease. The gene is coexpressed with one or more colon cancer
CC genes selected from carbonic anhydrase I, II and IV (CA I, II and IV),
CC carcinoembryonic antigen family of proteins (cea), colorectal carcinoma
CC tumour associated antigen (CO-029), down-regulated in adenoma (dra),
CC fatty-acid binding protein (fabp), galectin (galectin), glutathione
CC peroxidase (gpx2), guanylin (guan), cytokeratin 8 and 20 (ker 8 and 20),
CC cadherin (cadherin) or intestinal mucin (muc-2). The present polynucleotide
CC sequence is useful for diagnosing a disease or condition associated
CC with altered expression of the gene. The polynucleotide and encoded
CC polypeptides are useful for treating or preventing such diseases. They
CC are useful for diagnosing colon cancer, metastatic colon cancer, atrophic
CC gastritis, cholecystitis, Crohn's disease, irritable bowel syndrome and
CC ulcerative colitis.
XX
SQ Sequence 2616 BP; 887 A; 488 C; 515 G; 726 T; 0 other;
XX
Query Match 53.8%; Score 136.2; DB 21; Length 2616;
Best Local Similarity 71.1%; Pred. No. 4,4e-35;
Matches 180; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
OY 1 AACAAAGTGGTCCATCATGACACAGTGGGCGCCCTCTGAGCTCAAGAATAG 60
DB 740 aacaagaggggccattgtcttattatgcttgggaagagctgctgtagagcagtaa 799
OY 61 AGGACCTGCCAAATGACAGAGGTTTACAGACATATGCTTCAGTCAAGTTACAGACA 120
DB 800 tagagatgagcaagaataagaggaagcattatttgcagtgagcagcagaaca 859
OY 121 ATGGCTCATGATGCTTTGGGCGCCCTTCATCAGGAATGAGAGCTGCTCAGCCGT 180
DB 860 atggcctcatgattgtcttgggctctacacaggaatactgtctcccaagaat 919
OY 181 CCATCAGCTTGAGAGTAAGGATTAACTCCAGAACAGCCAGTGAATGACAGAG 240
DB 920 ccctcagctcgaagtaaggaataacacactgaatagtaatcctggaatgaacagacag 979
OY 241 TGATCGTGGACAG 253
DB 980 tcataatgtag 992
XX
RESULT 14
AAZ65095
ID AAZ65095 standard; cDNA; 3265 BP.
XX
AC AAZ65095;
XX
XX 05-APR-2000 (first entry)
DE Membrane-bound protein PRO1124 encoding cDNA.
XX
XX Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping; ss.
XX
XX Homo sapiens.
OS
XX
XX WO963088-A2.
XX
XX PD 09-DEC-1999.
XX
XX 02-JUN-1999; 99WO-US12252.
XX
XX

XX
XX 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088032.
PR 04-JUN-1998; 98US-0088036.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 06-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0088906.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089807.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 23-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR

XX	DR	P-PSDB; AA166749.	XX
XX	PI	Wood WJ, Yuan J;	XX
XX	PI	Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK,	XX
PA	(GETH) GENENTECH INC.		PA
PR	25-JUN-1998;	98US-0096936.	PR
PR	26-JUN-1998;	98US-0096862.	PR
PR	26-JUN-1998;	98US-0090863.	PR
PR	01-JUL-1998;	98US-0091358.	PR
PR	01-JUL-1998;	98US-0091360.	PR
PR	01-JUL-1998;	98US-0091544.	PR
PR	02-JUL-1998;	98US-0091478.	PR
PR	02-JUL-1998;	98US-0091519.	PR
PR	02-JUL-1998;	98US-0091626.	PR
PR	02-JUL-1998;	98US-0091628.	PR
PR	02-JUL-1998;	98US-0091633.	PR
PR	02-JUL-1998;	98US-0091646.	PR
PR	02-JUL-1998;	98US-0091673.	PR
PR	07-JUL-1998;	98US-0091978.	PR
PR	07-JUL-1998;	98US-0091982.	PR
PR	09-JUL-1998;	98US-0092182.	PR
PR	10-JUL-1998;	98US-0092477.	PR
PR	20-JUL-1998;	98US-0093339.	PR
PR	30-JUL-1998;	98US-0094651.	PR
PR	04-AUG-1998;	98US-0095282.	PR
PR	04-AUG-1998;	98US-0095285.	PR
PR	04-AUG-1998;	98US-0095301.	PR
PR	04-AUG-1998;	98US-0095302.	PR
PR	04-AUG-1998;	98US-0095310.	PR
PR	04-AUG-1998;	98US-0095321.	PR
PR	04-AUG-1998;	98US-0095325.	PR
PR	10-AUG-1998;	98US-0095916.	PR
PR	10-AUG-1998;	98US-0095929.	PR
PR	10-AUG-1998;	98US-0096012.	PR
PR	11-AUG-1998;	98US-0096143.	PR
PR	11-AUG-1998;	98US-0096146.	PR
PR	11-AUG-1998;	98US-0096329.	PR
PR	12-AUG-1998;	98US-0096357.	PR
PR	17-AUG-1998;	98US-0096756.	PR
PR	17-AUG-1998;	98US-0096768.	PR
PR	17-AUG-1998;	98US-0096773.	PR
PR	17-AUG-1998;	98US-0096791.	PR
PR	17-AUG-1998;	98US-0096867.	PR
PR	17-AUG-1998;	98US-0096891.	PR
PR	17-AUG-1998;	98US-0096894.	PR
PR	17-AUG-1998;	98US-0096895.	PR
PR	17-AUG-1998;	98US-0096897.	PR
PR	18-AUG-1998;	98US-0096949.	PR
PR	18-AUG-1998;	98US-0096950.	PR
PR	18-AUG-1998;	98US-0096959.	PR
PR	18-AUG-1998;	98US-0096960.	PR
PR	18-AUG-1998;	98US-0097022.	PR
PR	19-AUG-1998;	98US-0097141.	PR
PR	20-AUG-1998;	98US-0097218.	PR
PR	24-AUG-1998;	98US-0097661.	PR
PR	26-AUG-1998;	98US-0097951.	PR
PR	26-AUG-1998;	98US-0097952.	PR
PR	26-AUG-1998;	98US-0097954.	PR
PR	26-AUG-1998;	98US-0097955.	PR
PR	26-AUG-1998;	98US-0097971.	PR
PR	26-AUG-1998;	98US-0097974.	PR
PR	26-AUG-1998;	98US-0097978.	PR
PR	26-AUG-1998;	98US-0097979.	PR
PR	26-AUG-1998;	98US-0097986.	PR
PR	26-AUG-1998;	98US-0098014.	PR
PR	31-AUG-1998;	98US-0098525.	PR
PR	16-SEP-1998;	98US-0100634.	PR
PR	12-JAN-1999;	99US-0115565.	PR
XX			XX

PR 18-FEB-2000; 2000MO-US04342.
PR 22-FEB-2000; 2000MO-US04414.
PR 01-MAR-2000; 2000MO-US05601.
PR 03-MAR-2000; 2000US-0187202.
PR 25-APR-2000; 2000US-0199397.
PR 22-MAY-2000; 2000MO-US14042.
PR 05-JUN-2000; 2000US-0209832.

XX (GETH) GENENTECH INC.

XX Eaton DL, Filvaroff E, Gerltsen ME, Goddard A, Godowski PJ,
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI,
XX

DR WPI; 2001-183260/18.
P-PSDB; AAB87560.

XX Eighty four nucleic acids encoding PRO polypeptides, useful in
PT molecular biology, including use as hybridization probes, and in
PT chromosome and gene mapping.

XX Claim 2; Fig 69; 278pp; English.

CC The present sequence is the coding sequence for a human PRO polypeptide
CC (secreted and transmembrane). The PRO protein, and PRO agonists, PRO
CC antagonists or anti-PRO antibodies are useful for preparation of a
CC medicament useful in the treatment of a condition which is responsive to
CC the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
CC protein may also be employed as molecular weight markers for protein
CC electrophoresis. The PRO coding sequence has applications in molecular
CC biology, including use as hybridisation probes, and in chromosome and
CC gene mapping.

XX Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;

Query Match 53.8%; Score 136.2; DB 22; Length 3265;
Best Local Similarity 71.1%; Pred. No. 4,8e-35;
Matches 180; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

OY 1 AACAAAGTGTGCATCATCACACAGTCGCTTGGGCCCTCTGCAGCTCAAGAACTAG 60
DB 1304 aacaaagtgtggtccatgttcatttattgtcttgggaagctgtctgaagaagcagtaa 1363
OY 61 AGGAGCTGTCCAAATGACAGAGAGTTTACAGACATATGCTTCAGATCAAGTTTCAACA 120
DB 1364 tagaagtgagcaagaatacaagaaggaagcatttattgttcagatgaagctcagaca 1423
OY 121 ATGGCCTCATTTGATGCTTTGGGCCCTTTCATCAGAAATGAGCTGTCTCAGCGCT 180
DB 1424 atggcctcatgtatgcttcttgggctcttaacacagaataactgactctccagaagt 1483
OY 181 CCATCCAGCTTGAGATTAAGGATTACCTCCAGAACAGCAGTGAATGAATGATGACAG 240
DB 1484 ccatcagcttgagaaagtaaggaagtaaacactgataagtaagctgtgaagaaagcactg 1543
OY 241 TGATCGTGGACAG 253
DB 1544 tcaaatgtatag 1556

Search completed: April 4, 2002, 03:17:48
Job time: 60220 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 02:59:52 ; Search time 22700.8 seconds

(without alignments)
183.861 Million cell updates/sec

Title: US-09-049-696-8

Perfect score: 253
Sequence: 1 AACCAAGTGTGTCATC.....GGCACAAGTATCGTGACAG 253

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBankl: *
1: gb_ba: *
2: gb_hg: *
3: gb_in: *
4: gb_com: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_ov: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rnd: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rnd: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	253	100.0	2022	9	AK024970	AK024970 Homo sapi
2	253	100.0	2826	9	AF127036	AF127036 Homo sapi
3	253	100.0	3311	6	AX193489	AX193489 Sequence
4	253	100.0	3311	9	AF039400	AF039400 Homo sapi
5	185.8	73.4	3079	4	AF095584	AF095584 Sus scrofa
6	174.6	69.0	2937	10	AB017156	AB017156 Mus muscu
7	136.2	53.8	1895	9	AK000138	AK000138 Homo sapi
8	136.2	53.8	3204	9	AF127035	AF127035 Homo sapi
9	136.2	53.8	3221	9	AK000072	AK000072 Homo sapi
10	136.2	53.8	3265	6	AX092338	AX092338 Sequence
11	111.6	44.1	35278	9	AF039401	AF039401 Homo sapi
12	111.6	44.1	113764	9	HS7551E10	HS7551E10 Human DNA
13	111.6	44.1	164891	2	AL358950	AL358950 Homo sapi
14	93.4	36.9	2832	9	AF127980	AF127980 Homo sapi
15	93.4	36.9	2970	9	AF043977	AF043977 Homo sapi
16	93.4	36.9	3604	9	AB026833	AB026833 Homo sapi
17	93.4	36.9	4077	6	AX054697	AX054697 Sequence
18	87	34.4	2820	4	AF001263	AF001263 Bos tauru
19	87	34.4	3288	4	AF001262	AF001262 Bos tauru
20	87	34.4	3317	4	AF001261	AF001261 Bos tauru
21	86.8	34.3	3058	10	AF108501	AF108501 Mus muscu
22	85.4	33.8	2984	4	BT036445	BT036445 Bos tauru
23	85.2	33.7	2765	10	AF115852	AF115852 Mus muscu
24	85.2	33.7	3137	10	BC008147	BC008147 Mus muscu
25	82	32.4	3022	10	AF047838	AF047838 Mus muscu
26	82	32.4	3471	10	AF052746	AF052746 Mus muscu
27	71.6	28.3	140718	2	AL356270	AL356270 Homo sapi
28	57.4	22.7	3415	9	AF043976	AF043976 Homo sapi
29	55	21.7	140718	2	AL356270	AL356270 Homo sapi
30	55	21.7	175591	2	AC068071	AC068071 Homo sapi
31	46.6	18.4	124096	2	AL354988	AL354988 Homo sapi
32	36.4	14.4	299550	1	AP001511	AP001511 Bacillus
33	35.6	14.1	163253	2	AL162575	AL162575 Homo sapi
34	35.6	14.1	187847	2	AC023155	AC023155 Homo sapi
35	35.6	14.1	199768	2	AL592490	AL592490 Homo sapi
36	35	13.8	69126	9	AL354752	AL354752 Human DNA
37	34.2	13.5	206506	2	AC025585	AC025585 Mus muscu
38	34	13.4	72417	2	AC016988	AC016988 Homo sapi
39	34	13.4	16167	9	AC005483	AC005483 Homo sapi
40	34	13.4	190569	2	AC084164	AC084164 Mus muscu
41	33.8	13.4	6533	14	AF293665	AF293665 Hantaviru
42	33.8	13.4	6533	14	AF336826	AF336826 Hantaviru
43	33.8	13.4	156604	9	AL356865	AL356865 Human DNA
44	33.6	13.3	35412	8	SPAC4E8	Z98530 S. Pombe chr
45	33	13.0	392	11	G61053	G61053 SHGC-85269

ALIGNMENTS

RESULT: 1
AK024970
LOCUS: AK024970
DEFINITION: Homo sapiens cDNA: FLJ21317 fls, clone COL02275, highly similar to AF127036 Homo sapiens calcium-activated chloride protein 1 (CACCL1) mRNA.
ACCESSION: AK024970
VERSION: AK024970.1 GI:10437397
KEYWORDS: oligo capsins; fls (full insert sequence).
SOURCE: Homo sapiens colon cDNA to mRNA, clone lib:COL clone:COL02275.
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE: 1 (sites)
Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S. NED human cDNA sequencing project
Unpublished (2000)
JOURNAL: 2 (bases 1 to 2022)

AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.

TITLE Direct Submission

JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: ednal@ms.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source

1. .2022

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="COL2275"

/tissue_type="colon"

/note="cloning vector pME18SFL3"

1. .2022

/note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CaCC1) mRNA"

BASE COUNT 612 a 472 c 453 g 485 t

ORIGIN

Query Match 100.0%; Score 253; DB 9; Length 2022;
Best Local Similarity 100.0%; Pred. No. 9.6e-74;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAAAGTGTGTCATCCACAGAGTCGCTTTGGGCGCTCTGCACTCAAGAACTAG 60
|||||
Db 490 AACAAAGTGTGTCATCCACAGAGTCGCTTTGGGCGCTCTGCACTCAAGAACTAG 549
|||||

QY 61 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAACA 120
|||||

Db 550 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAACA 609
|||||

QY 121 ATGGCTCATTTGATGCTTTTGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCGCT 180
|||||

Db 610 ATGGCTCATTTGATGCTTTTGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCGCT 669
|||||

QY 181 CCATCCAGCTTGAAGTAAAGGATTAACCTCCAGAACAGCCAGTGAATGGACAG 240
|||||

Db 670 CCATCCAGCTTGAAGTAAAGGATTAACCTCCAGAACAGCCAGTGAATGGACAG 729
|||||

QY 241 TGATCGTGGACAG 253
|||||

Db 730 TGATCGTGGACAG 742

RESULT 2

AF127036 2826 bp mRNA PRI 10-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)
DEFINITION mRNA, complete cds.
ACCESSION AF127036
VERSION AF127036.1 GI:4585468
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 2826)
AUTHORS Agnel, M., Vernat, T. and Culouscou, J.-M.
TITLE Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea

JOURNAL FEBS Lett. 455 (3), 295-301 (1999)

MEDLINE 99364503

PUBMED 10437792

REFERENCE 2 (bases 1 to 2826)

AUTHORS Agnel, M. and Culouscou, J.-M.

TITLE Direct Submission

JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrières, Rueil-Malmaison 92500, France

FEATURES

source

1. .2826

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="small intestine; colon"

1. .2826

/gene="CaCC1"

5. .2749

/gene="CaCC1"

/note="bovine epithelial chloride channel homolog"

/product="calcium-activated chloride channel protein 1"

/protein_id="A4D25487.1"

/db_xref="GI:4585469"

/translation="MGPRKSSVPIILHLLEGALNSLHOLNNNGEIVAIIDPNP
EDPTLQIQKQWVTSIFRATGRRFRKRVAVILIPETWTKADYVPRKLEYKNA
DVLAVESTIPGNDERTTEOMNGEKEGERIHLTPPTAKRTAETGPOGRARVHEMAH
LWGEFDEYNNDKEFTLSNGRIQAVNCSAGITGVNKKCGGSCYTKCTNKYTG
YKGECEVLOSROTEKASIMFAOHVDISIEFCEVDHNNKSNKONKONLSTWETVL
RSEDEKRTTPMTQPPNPFSLLIQGRIVCLVDKSGSMATGNLNLNAGOLF
LOTVELGSWGWGTFDPAHVOSELQIINSRSDRLRLPAASGCTSGISLSA
FYIRKKRPPDGEIYLTDGEENTISGCFNEVSGKALIHVVALPSAOLELEISK
MGGLQTVASDQVONNGILIDARGLSSGNVAVSORSIOLESGLTLONSOMNGVIV
DSTVGKDTLEFLITWTQPPQIILMDPSGKOGSGFYVDKTKMAVLOIPILAVGTMKY
SIALSSQTLIVTSTRASNAVLPITVTSKINDKSPSPLYVYANINQGSPIIRA
SVALLSEVNVKRVTLLELDNGACATADDDVYSKFTYDINGTSKVYALGAVN
AARRVYIPQSGALYIPGWIENDEIQMNPBPINDDVDVQHVCFSRSGSSEVVA
DVNPAPLPDLPFGQITDLKAEIHGSLINLTETAPDGDHGTAKHYIRISTSLD
LRKFNSLOVNTLILPEANSEVEFLPEPENTIPENTGTDLEIAIOAVDKVLSKEI
SNARVSLRIPPTPPTPPTSPDTSAPCPNINHNSIIPGHIHKIMKWTGELQSLIA

BASE COUNT 875 a 623 c 632 g 696 t

ORIGIN

Query Match 100.0%; Score 253; DB 9; Length 2826;
Best Local Similarity 100.0%; Pred. No. 1e-73;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAAAGTGTGTCATCCACAGAGTCGCTTTGGGCGCTCTGCACTCAAGAACTAG 60
|||||

Db 1231 AACAAAGTGTGTCATCCACAGAGTCGCTTTGGGCGCTCTGCACTCAAGAACTAG 1340
|||||

QY 61 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAACA 120
|||||

Db 1341 AGGAGCTGTCCAAATGACAGAGGTTTACAGACATATGCTTCAGATCAAGTTCAACA 1400
|||||

QY 121 ATGGCTCATTTGATGCTTTTGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCGCT 180
|||||

Db 1401 ATGGCTCATTTGATGCTTTTGGGGCCCTTTCATCAGAAATGAGCTGTCTCTCAGCGCT 1460
|||||

QY 181 CCATCCAGCTTGAAGTAAAGGATTAACCTCCAGAACAGCCAGTGAATGGACAG 240
|||||

Db 1461 CCATCCAGCTTGAAGTAAAGGATTAACCTCCAGAACAGCCAGTGAATGGACAG 1520
|||||

QY 241 TGATCGTGGACAG 253
|||||

Db 1521 TGATCGTGGACAG 1533

RESULT 3

AX193489 3311 bp DNA PAT 15-AUG-2001
LOCUS AX193489
DEFINITION Sequence 1056 from Patent WO0149716.
ACCESSION AX193489

VERSION AX193489.1 GI:15211440
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: WO 0149716-A 1056 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 100.0%; Score 253; DB 6; Length 3311;
Best Local Similarity 100.0%; Pred. No. 1e-73;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AACAAAGTGGTCCATCATCCACAGAGTGGGGCCCTGCGAGCTCAAGAACTAG 60
|||||
DB 1628 AACAAAGTGGTCCATCATCCACAGAGTGGGGCCCTGCGAGCTCAAGAACTAG 1687
|||||
OY 61 AGGAGCTGTCACAAATGACAGAGGTTTACACATATGCTTCAGATCAAGTTCAAGACA 120
|||||
DB 1688 AGGAGCTGTCACAAATGACAGAGGTTTACACATATGCTTCAGATCAAGTTCAAGACA 1747
|||||
OY 121 ATGGCTCATTTGATGCTTTGGGGCCCTTTCATCAGAAATGAGTGTCTCAGCGCT 180
|||||
DB 1748 ATGGCTCATTTGATGCTTTGGGGCCCTTTCATCAGAAATGAGTGTCTCAGCGCT 1807
|||||
OY 181 CCATCAGCTTGAGAGTAAGGATTAACCTCCAGAACAGCCAGTGTGATGAATGGACAG 240
|||||
DB 1808 CCATCAGCTTGAGAGTAAGGATTAACCTCCAGAACAGCCAGTGTGATGAATGGACAG 1867
|||||
OY 241 TGATCGTGGACAG 253
|||||
DB 1868 TGATCGTGGACAG 1880
|||||

RESULT 4
AF039400 3311 bp mRNA PRI 14-DEC-1998
LOCUS Homo sapiens calcium-dependent chloride channel-1 (hCICL1) mRNA,
DEFINITION complete cds.
ACCESSION AF039400
VERSION AF039400.1 GI:4009457
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Gruber,A.D., Elble,R.C., Ji,H.L., Schreier,K.D., Fuller,C.M. and
Pauli,B.U.
TITLE Genomic cloning, molecular characterization, and functional
analysis of human CICL1, the first human member of the family of
Ca²⁺-activated Cl⁻ channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)
MEDLINE 99047526
REFERENCE 2 (bases 1 to 3311)
AUTHORS Gruber,A.D., Elble,R. and Pauli,B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
Cornell University, Ithaca, NY 14853, USA
FEATURES
source Location/Qualifiers
1..3311

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/tissue_type="small intestine"
1..3311
/gene="hCICL1"
352..3096
/gene="hCICL1"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPEKSVFILLHLEGALSNSLIQLNNNGEIVADPNP
DETLIOQKDVATQASLYFEATGRKRFKRVAVALLIPTTKADYRPAKLEYKNA
DYLVAESPPGDEPYTBMGMCGRKGRVILHTPPFIAGKLAEGPGOKAVVEMAH
LWGVFDEYNDEKFTYLSNGRLOAVRCSAGITGVNVVKRCQGGSCYTKRCKFTWGL
YKGEFEVYLSQKOTERASIMFAQHVDSIVFECTEQNHNEKAPNOKNOKNLSSTVEI
RDEDEFKTTPMTQTPNPFTSLDQIRIVCLVDKSGSNATGRNLRLNAGOLF
LOFVELGSWGVATFDSAHVQSELIQINSDEGRDRLAKRLPAASGGTSIGLSRA
FTVIRKVPDQSEIVLITDGEDNTISGFENVRGSAIITHTVALGPSAOGLELSK
MGGQTATASDVQVONNGLDARGALSSNGANVSQSIQESGLTLQNSQNMNTYIV
DSTVGKDTLFLTWTPQPIILMDPSGQKGFVVDKMTKNAVLIQIPELAVGTWKY
SIQASSQTLTIVTSTRASNATLPITVTSKTKKDYSPSPVLVAVANIRQASPLRA
SVTALIESVNGKTVTLELDNGACADATDDDVYSRFTYDTNGRSVAVAGLVN
AARRVIRPOSGALYIPGWIENDEIOMNPDEIKNDVOYHKOVCESRTSSGSPVAS
DVNPAIDPLPFGIOTDKARIEHGSILNLTWTAAGDDYDHGTAHKYIRITSLTD
LKDKNESQVVTATLALIKRENSSEVFLFKPNITPENTDIFALQAVDKDLKLSAI
SNIAVSLFIPQTPETPSPDETAPCPNINHSTIREHILKIMKWTGELQLSIA

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 100.0%; Score 253; DB 9; Length 3311;
Best Local Similarity 100.0%; Pred. No. 1e-73;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AACAAAGTGGTCCATCATCCACAGAGTGGGGCCCTGCGAGCTCAAGAACTAG 60
|||||
DB 1628 AACAAAGTGGTCCATCATCCACAGAGTGGGGCCCTGCGAGCTCAAGAACTAG 1687
|||||
OY 61 AGGAGCTGTCACAAATGACAGAGGTTTACACATATGCTTCAGATCAAGTTCAAGACA 120
|||||
DB 1688 AGGAGCTGTCACAAATGACAGAGGTTTACACATATGCTTCAGATCAAGTTCAAGACA 1747
|||||
OY 121 ATGGCTCATTTGATGCTTTGGGGCCCTTTCATCAGAAATGAGTGTCTCAGCGCT 180
|||||
DB 1748 ATGGCTCATTTGATGCTTTGGGGCCCTTTCATCAGAAATGAGTGTCTCAGCGCT 1807
|||||
OY 181 CCATCAGCTTGAGAGTAAGGATTAACCTCCAGAACAGCCAGTGTGATGAATGGACAG 240
|||||
DB 1808 CCATCAGCTTGAGAGTAAGGATTAACCTCCAGAACAGCCAGTGTGATGAATGGACAG 1867
|||||
OY 241 TGATCGTGGACAG 253
|||||
DB 1868 TGATCGTGGACAG 1880
|||||

RESULT 5
AF095584 3079 bp mRNA MAM 13-OCT-2000
LOCUS Sus scrofa epithelial chloride channel protein (AEC) mRNA,
DEFINITION complete cds.
ACCESSION AF095584
VERSION AF095584.1 GI:6002645
SOURCE pig.
ORGANISM Sus scrofa
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 3079)

Db	1414	AVGCGCTTGTTGATGCTTTCCGACGACGACGCTCTCTCAGGAANAATCGGCGCATTCGCTCAGCACT	1473
Oy	131	CCATCCAGCTTGAGAGCTTAAGGATTAACCCCTTCAGACAACCGATGATGAATGCGACAG	240
Db	1474	CCATCCAGCTTGAGAGCGGGAGGTAACTTCACGAATAACCAATGATGAATGCGCTCAG	1533
Oy	241	TGATGTCGACAG	253
Db	1534	TGATGTCGACAG	1546
RESULT	7		
AK000138			
LOCUS	AK000138	1895 bp	mRNA
DEFINITION	Homo sapiens cDNA FLJ20131 fls, clone COL06357.		22-FEB-2000
ACCESSION	AK000138		
VERSION	AK000138.1	GI:7020030	
KEYWORDS	oligo capping; fls (full insert sequence).		
SOURCE	Homo sapiens colon cDNA to mRNA, clone_11b; COL clone:COL06357.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Ota, T., Suzuki, Y., Obyashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.		
TITLE	NEBO human cDNA sequencing project		
JOURNAL	Unpublished (2000)		
REFERENCE	2 (bases 1 to 1895)		
AUTHORS	Sugano, S., Suzuki, Y., Ota, T., Obyashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.		
TITLE	Direct Submision		
JOURNAL	Submitted (15-FEB-2000) to the DDBJ/EMBL/Genbank databases. Sumio Sugano, Institute of Medical science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)		
COMMENT	NEBO human cDNA sequencing project supported by ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction. 5'-6'3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).		
FEATURES			
source	Location/Qualifiers		
	1..1895		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="COL06357"		
	/clone_11b="COL"		
	/tissue_type="colon"		
	/note="cloning vector pME18SFL3"		
	66..1475		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="BAA90969.1"		
	/db_xref="GI:7020031"		
	/transcript="MSKRTGSHFFYSDEAONNGILDFGALTSGMTLSOXLSOLEE		
	KGLINSLNANNDVILIDSTYKDFPFLITINSIPSTISMDPGTILMENTVATSTK		
	MAYLSIETGAKVGTMAINLQAKAPETITVTYSRANSSVPTTVAKAKMDVNSFF		
	SPMLVVEILOGIVPVLGANTATIESONGTVEVLELIDNGAGADSEKNDVSRYTF		
	ATYDNGRSILKVRAGAGANTARLRLPLNAAAYIPGVVNGELIENAPPEIDEDT		
	TTEDEFSRTASGGAFFVQVSLPLPDQPPSQITDDAVYHEDKILITWAPDNGF		
	VKGQRTIIRISASILDRSGFDALQVNTDLSPKREANSRSPAFKRENTSENATTH		
	IFVIRKIDRSNLTSTKYSNTAQTAVLFIQANPDIDIPPTPTTPIDKSHNSGVNIST		
	VLSTIGSVIVYVNFILSTTI"		
BASE COUNT	658 a 352 c 357 g 528 t		
ORIGIN			
Query Match	53.8%; Score 136.2; D: 9; length 1895;		
Best Local Similarity	71.1%; Pred. No. 1.4e-34;		
Matches 180; Conservative	0; Mismatches 73; Indels 0; Gaps 0;		

[illegible]

Db 24575 GCTTTGGGGCCCTTCATCAGAAATGAGCTGTCTCTCAGCGCTCCATCCAGGTTG 24632

RESULT 12

LOCUS

HSJ651E10 113764 bp DNA PRI 27-MAY-2000
Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1,
complete sequence.

DEFINITION

ACCESSION

AL122002

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Jun 4, 2000 this sequence version replaced g1:8247030.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
EMBL, SWISSPROT, TrEMBL, WP, WormPeP, Information
on the WormPeP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP4-651E10 is from the library RPCR-4 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP4-651E10.

FEATURES

Source

Location/Qualifiers
1. 113764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p22.3-31.1"
/clone="RP4-651E10"
/clone_1lb="RPCR-4"
BASE COUNT 34890 a 21989 c 22351 g 34534 t
ORIGIN

Query Match

Best Local Similarity 44.1%; Score 111.6; DB 9; Length 113764;
Matches 114; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 75 ATGACAGAGGTTACAGACATATGCTTCAGATCAAGTTGAGCAATGGCCCTCATTTGAT 134

Db 24228 ATTATAGAGGTTTACAGACATATGCTTCAGATCAAGTTGAGCAATGGCCCTCATTTGAT 24287

QY 135 GCTTTGGGGCCCTTCATCAGAAATGAGCTGTCTCTCAGCGCTCCATCCAGGTTG 192

Db 24288 GCTTTGGGGCCCTTCATCAGAAATGAGCTGTCTCTCAGCGCTCCATCCAGGTTG 24345

RESULT 113

ALJ38950/c

LOCUS

DEFINITION

ACCESSION

ALJ38950

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

ALJ38950 164891 bp DNA HTG 23-JAN-2001
Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN
PROGRESS ***; 10 unordered pieces.
ALJ38950.4 GI:12539689
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 164891)
Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jan 26, 2001 this sequence version replaced g1:9988471.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: BA444C12
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least Q40
Consensus quality: 162610 bases at least Q30
Consensus quality: 163328 bases at least Q20
Insert size: 163991; sum-of-contigs
Insert size: 163373; 6.6% error; agarose-IP
Quality coverage: 5.19x in Q20 bases; sum-of-contigs quality
coverage: 5.21x in Q20 bases; agarose-IP

NOTE: This is a 'working draft' sequence. It currently
consists of 10 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 34236: contig of 34236 bp in length
* 34237 34336: gap of 100 bp
* 34337 47372: contig of 13036 bp in length
* 47373 47472: gap of 100 bp
* 47473 67025: contig of 19553 bp in length
* 67026 67125: gap of 100 bp
* 67126 71239: contig of 4114 bp in length
* 71240 71339: gap of 100 bp
* 71340 83903: contig of 12564 bp in length
* 83904 84003: gap of 100 bp
* 84004 91357: contig of 7354 bp in length
* 91358 91457: gap of 100 bp
* 91458 96452: contig of 4995 bp in length
* 96453 96552: gap of 100 bp
* 96553 148189: contig of 51637 bp in length
* 148190 148289: gap of 100 bp
* 148290 159778: contig of 11489 bp in length
* 159779 159878: gap of 100 bp
* 159879 164891: contig of 5013 bp in length.

FEATURES

Source

Location/Qualifiers
1. 164891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-444C12"
/clone_1lb="RPCR-11.2"
misc_feature 1. 34236

/note="assembly-fragment:00027
fragment_chain:1
clone_end:77
vector_side:left"
misc-feature 34337..47372
/note="assembly-fragment:00312
fragment_chain:1"
misc-feature 47473..67025
/note="assembly-fragment:01923
fragment_chain:1"
misc-feature 67126..71239
/note="assembly-fragment:02094
fragment_chain:1"
misc-feature 71340..83903
/note="assembly-fragment:02213
fragment_chain:1"
misc-feature 84004..91357
/note="assembly-fragment:01292
fragment_chain:1"
misc-feature 91458..96452
/note="assembly-fragment:01391
fragment_chain:1"
misc-feature 96553..148189
/note="assembly-fragment:02045
fragment_chain:1"
misc-feature 148290..159778
/note="assembly-fragment:00427
fragment_chain:1"
misc-feature 159879..164891
/note="assembly-fragment:00773
fragment_chain:1"
vector_side:right"
BASE COUNT 4881 a 32179 c 32357 g 51067 t 907 others
ORIGIN

Query Match 44.1%; Score 111.6; DB 2; Length 164891;
Best Local Similarity 96.6%; Pred. No. 4.9e-26;
Matches 114; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 75 ATGACGAGGAGTTTACAGACATATGCTTCAGATCAAGTTGAGCAACATGGCCCTCATTTGAT 134
DB 57487 ATTTTAGGAGGTTTACAGACATATGCTTCAGATCAAGTTGAGCAACATGGCCCTCATTTGAT 57428
QY 135 GCTTTGGGCGCCCTTCATCAGAAATGAGCTGTCTCTCAGCGCTCCATCCAGCTTG 192
DB 57427 GCTTTGGGCGCCCTTCATCAGAAATGAGCTGTCTCTCAGCGCTCCATCCAGCTTG 57370
RESULT 14
AF127980 2832 bp mRNA PRI 11-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 3 (CAC3)
DEFINITION MRNA, complete cds.
ACCESSION AF127980
VERSION AF127980.1 GI:5726286
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2832)
AUTHORS Agnel,M., Vernet,T., and Culouscou,J.M.
TITLE Identification of three novel members of the calcium-dependent
chloride channel (CAC3) family predominantly expressed in the
digestive tract and trachea
JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792
REFERENCE 2 (bases 1 to 2832)
AUTHORS Agnel,M. and Culouscou,J.-M.
TITLE Direct Submission

JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des
Carrières, Rueil-Malmaison 92500, France
FEATURES
Location/Qualifiers
source 1..2832
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1..2832
/gene="CAC3"
CDS 1..2832
/gene="CAC3"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 3"
/protein_id="A4D8397.1"
/db_xref="gi:5726287"
/translation="MTORSIAGPICNLRKFTVLIVLISLSELPFAGVVOLODNGNGL
TAINPVPENOLISIKKMTIBASYLELNATKRRFRFNKILIPATYKANNSEIK
QSEYKANIIVTDWGAHDDPTLQYRCGRBKGYTHFPNPLINDNLPAIGYSGR
VPEVHEAHLRMGFDEYENDKPEYINGQIKVTRCSSDITGIFVCEKGPQENCI
SKLKEGCTFEINSONAVSIMFQSLSSVEFCASVHNOEAPMLQMCSLRSAM
DYITDSADPHSHSPMNGTELPEPTFSILOAGGVYCLVDVSKMAEDRLQLQQA
AEFLYMOIVEIHTEFVGLASFDSKGEIRAOHINSNDKRLVSYLPTVSAKTQISI
CSGLKGFVEYELNKGAGSVMIIVTSDDKILGCLPTVLSGSHISIALGSSAA
PNLELSRLNGLKFPEVPDISNSMIDAFSRISSTGDIPOOHIOLESTGEVAPFH
OLKNTVTVNTVGNDMPLVTWQASGPPEILFDPGRVYFVTFNLTNLPTFASLMI
PGTAKPGHWYTLNMTNHSLOALKTVTSRASNSAPVATVEAFVSRDSLHPFPYMI
YANVKGCFEPLMAVATVATVEPTGDPVTRLLDDAGVDALINDISRYFSPYAN
GRYSLVHNHNSPISITSPAHISIPGSAMVATVANGNOMNAPRSVGRNEERKAG
FSRVSSGGSFVYGVAPGHPVFPCKIIDEAVYVEBELISVAPREDPDQCAT
STRIRMSKSLQNTODFPNNAITLVNTSIRNPQAGIREITFFSPQISTNGPEQPNET
HSHRIYVAIRADRNLSIOSAVSNIAQAPLFIIPNSDPVARDYLILKGVLTAMGLIG
IICLIIVVAIRHHTLSRRKRADKRENGTKLL"
BASE COUNT 877 a 609 c 594 g 752 t
ORIGIN

Query Match 36.9%; Score 93.4; DB 9; Length 2832;
Best Local Similarity 61.1%; Pred. No. 3.4e-20;
Matches 151; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 6 ATGTGTCATCATCCACACAGTCGCTTGGGCGCCCTCTCAGCTCAAGAACATGAGAG 65
DB 1306 AGTGGTTCAACAAATTCACCTCATTTGCCCTGGGTTCATCTCAGCCCAAAATCTGAGAGAA 1365
QY 56 CTGTCCAAATGACAGAGAGTTTACAGACATATGCTTCAGATCAAGTTGAGCAACATGGC 125
DB 1356 TTATACAGCTTTACAGAGAGTTTAAAGTTCTTGTTCACATATATCAAACTCCAAATAGC 1425
QY 126 CTCATTGATGCTTTTGGGCGCCCTTCATCAGAAATGAGCTGTCTCTCAGCGCTCCATC 185
DB 1426 ATGATTGATGCTTTCAATGAATTTCTCTGAACTGGAACATTTTCCAGCAACATATTT 1485
QY 186 CAGCTTGAAGTAAGGATTAACCTCCAGAACAGCAAGTGAATGGCAGACATATC 245
DB 1486 CAGCTTGAAGTAAGGATTAACCTCCAGAACAGCAAGTGAATGGCAGACATATC 1545
QY 246 GTGGACA 252
DB 1546 GTGGATA 1552
RESULT 15
AF043977 2970 bp mRNA PRI 23-JUN-1999
LOCUS AF043977
DEFINITION Homo sapiens calcium-activated chloride channel-2 (hCLCA2) mRNA,
complete cds.
ACCESSION AF043977
VERSION AF043977.1 GI:5138899
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

EDVALLIHIKIMYQOASPYLEATEAGTFFEFVNAILLIPESKAKPEPTRLPELEFFKNA
 DYALSTSPLENDPIETHEIGACGEGKRIHLIPDFDLAKGLIYQDPDRFEVEMAH
 FRMGVFNENNDKEFTYLSKGPQAVRSGALDQNOVRGCGSCITNGKQIDVING
 LYKNCVFPVDPHONERKASIMFNENNTISVYEFCTEKNHNDQAPNDQRCULBESTEY
 IOEEDERKOTTPTMTQCPAPFPFSLQIGORVCGIYLDNSGMLDNDRIINRNQASRLLE
 LLQTVGEGSGWAGVITFDSAAVYSELKQNLNGARGDRLIKHLPVSAQGSITCSGLTR
 AFATYIKRRYPSIDSEIVLADGEDNTIISCCDLVAKOSGALLHTVALGPAALKEQLS
 KMTGGLQTSDDQVQNNGLVDAPFALSSGNMAIKQHSIQTLESRCVNLQNNOMNGSVY
 VDSVGKQOTLELITWTHPTPIFTIMDSGVGONCEFIIDTTPKVAIYQPGTAKGFEMW
 YSIQASSOTLELUTVTSRASATILPPIYTPPVNKNGTGKFPSPVNVASIRGASPDILH
 ASVALLESVNGKTVTELELDNAGACATKNDVYSGRFPFQDNNGRSYXIMALGQ
 TSDQKRAAPRKNRPAYIDGMIEDEGVRRNPNRPRTSYQDQQLCFSTRISGGGSYVART
 VPAAPLPDLEPPCQIITDLKASIQGQNLVNLITWAPGDEYDHGASATIMKTSIYVLE
 LRDHFNESLQVNTTGLIPKENSSEIEFEELGZFGNGDIFALIAVDKSNLKSSE
 SNLARVSVPFIPOADPPIPEDSTPPCPDISINSTIPGJHVLIKIMKMWLGEOVTGLH
 2912: 2917
 polyA_signal
 polyA__signal
 polyA_site
 2933
 /note="putative"
 /note="putative"
 BASE COUNT 845 a 722 c 697 g 669 t
 ORIGIN

Query Match	Best Local Similarity	Score	Pos. 148-8	DB 12	Length 2933
Matches 183	Conservative	0	Mismatches 57	Indels	0
0%	76.2%	92.0%	0	0	0

FEATURES	SOURCE
1	AAAAAGCTGATGTTCTGGTTCGTCAGCTCTACCTCCAGAGTAATGATGAACCTACACTG 60
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	AGCAGATGGGCACACTGCGAGAGAGAAGGGGTGAAGAGATCCACCTCACTCTGATTTTCATG 120
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	
101	
102	
103	
104	
105	
106	
107	
108	
109	
110	
111	
112	
113	
114	
115	
116	
117	
118	
119	
120	
121	CAGGAAAAAATAGTACGTATATGACCAACAGGTGGCCATTGTTCATGATGGGCTC 180
122	
123	
124	
125	
126	
127	
128	
129	
130	
131	
132	
133	
134	
135	
136	
137	
138	
139	
140	
141	
142	
143	
144	
145	
146	
147	
148	
149	
150	
151	
152	
153	
154	
155	
156	
157	
158	
159	
160	
161	

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image:41c2662"
/clone_1lb="NCI_CGAP_L19"
/lab_host="DH10B (TI phage-resistant)"
/organ="liver; Vector: pCMV-SPORT6; site_1: NotI;
site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT
269 a 223 c 204 g 180 t
ORIGIN
1

```

```

Dy      1 AAATGCGTGTCTTGTTGGTTGCTAGTACTCCTCCACGTAATTGATAAACCCTCACTG   60
Db      3.11 AAAAGCGTAGTCCTTATTATCACAAACACGAGGCCCTTAGGAATGATGAGGCCTACACCG   370
Dy      .61 AGCATATGGGCAACTGTGTGAGAGAAGGGGTGAAAAGGATCCACCTCATCTGATTTCATTTG   120
Db      3.71 |---|-----|-----|-----|-----|-----|-----|-----|    430
          AACATATTAGAACCATGTGTGAGAAAAAAGGGGATTCAGGATTCACCTGACTGCTGACTTCTTGA
Dy      1.21 CAGGAAAAAAGTTTACCTGAATATGAGACCAACAAGTAGGGCATTTGTCCATGAGTGGGCTC   180
Db      4.21 CGGAAACAAAGCTGACTCAGTATGTGGCCACAAAGACAGACGACCTTTGTCCATGAGTGGGCTC   490
Dy      1.21 ATCTACGATGGGAGATATTTTGCACGAGTCACAAATATGATGAGAAAAATTCTACTATTCATG   240
Db      4.21 ACTTCGATGGGAGAGTGTCTAATGTAATACAAACAGACGAGAAGTTTCATTATCCAAG   550

RESULT           8
AL602665        600 bp       mRNA               EST             14-Aug-2001
LOCUS           DKFZPB8600214.t1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION      DKFZPB8600214 5', mRNA sequence.
ACCESSION       AL602665
VERSION         AL602665
KEYWORDS        EST.
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE       1 (bases 1 to 600)
AUTHORS        Mambutt,R., Heubner,D., Mewes,W., Well,B. and Wiemann,S.
TITLE          EST (Mambutt,R., Heubner,D., Mewes,H.W., Well,B. and Wiemann,S.)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Mambutt R
                MIPS
Am Klopfersplitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ). Email s.wiemann@dkfz-heidelberg.de;
sequenced by ACOMA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl.sequence available.
This Clone (DKFZPB8600214) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
Source
Location/Qualifiers
1..600
/oranism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZPB8600214"
/clone_lib="686 (synonym: hlcc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/vector="Vector: pTRipleX2; Site_1: SfIA; Site_2: SfIIB;
```


Accession	Sequence	Position
Db	385 AACATATAGGACGACATGTGGAGAAACAGGGGATTCAGATTCACCTGACTCTGACTTCTTA	444
OY	120 GCAGCA-AAAAAGTTAGCTGAATATGACACACCAAGTAGGGCATTTGTCATGATGCGGC	178
Db	445 GCAGGACAAAGAGCTGACACTAGTATGGGCGCAACAGACAGGACTTTGTTCATGATGCGGC	504
OY	179 TCATCTACGATGGGAGTATTT	200
Db	505 TCACCTCCGATGGGAGACTTTT	526

RESULT						
11						
BE382110	BE382110	522 bp	mRNA	EST	21-JUL-2000	
LOCUS	60127227270F1	NCI-CGAP_Mam1	Mus musculus	cDNA clone	IMAGE:3601672	5'
DEFINITION	mRNA sequence.					
ACCESSION	BE382110					
VERSION	BE382110.1	GI:9327475				
KEYWORDS	EST.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 522)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC).
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLAM8787 row: k column: 17
high quality sequence skip: 522.

FEATURES	Location/Qualifiers
source	1. .522

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3601672"
/clone_1ib="NCI_CGAP_Maml"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; site_1: SalI;
site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

```

Query Match	39.48;	Score 94.6;	DB 10;	Length 522;
Best Local Similarity	62.48;	Pred. No. 4.5e-18;		
Matches 148;	Conservative 0;	Mismatches 89;	Indels 0;	Gaps 0

OY	1	AAAAAGTCGATGTTCTGGTTGCTGAGTCTACTCTCCAGGTAATGATGAACCTTAACCTG	60
Db	105	ACAAAGCAGAGCGTCATAGTTGGCGATCCCTACACTTGAACATGGAGACGACCCCTACACCC	164
OY	61	AGCAGATGGGCACTGTGTGAGAGAAAGGTGAAGAATCCACTCACTCCTGATTTCATTG	120
Db	165	TTCACTATGAGCAAGTGTGGGACACAGGACACATACATCACTCACTCAAACTTCTCAG	224
OY	121	CAGGAAAAAGTTACTGAATGTGAGACCAACAGTAAAGCCATTTGTCCATGAGTGGGCTC	180
Db	225	TCACGTAACTACTCGCTATCTATGGACCCCGAGAGAGATCTTTGTCCATGAGTGGGCCC	284

Qy 181 ATCTACGATGGGAGTATTTCACGATGCATAATGTACGAATTCTACTATCCA 23
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dp 285 ATCTCCGGTGCGAGCTATTTTGATGATGATTAACGTGACGCCGCCCTTTCTACATTTCTCA 341

LOCUS	538 bp	EST	05-JUL-2001
DEFINITION	B1156386 mRNA		
	602919638f1 NIH_CGAP_Mam3	Mus musculus	cdna clone
	mRNA sequence.		
		IMAGE:5060097	5'

ACCESSION	BI156386
VERSION	BI156386.1
KEYWORD	EST
SOURCE	house.mouse
ORGANISM	Mus musculus

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Eumariyola; Eutacoa; Choridata; Craniata; Euteleostomi.
 Mammalia; Neohelina; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 538)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.

Email: c9apbs@email.nih.gov
Tissue Procurement: Lochar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLAM1162 row: k column: 10
High quality sequence stop: 356.

FEATURES	Location/Qualifiers
source	1. .538

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5060097"
/clone_lib="NIH_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: PCMV-SPORT6; Site:1: NotI;
Site_2: SalI; Cloned unidirectionally, primer: Oligo dr.
Average insert 2 kb. Library constructed by Life
Technologies, catalog #12017-018. Investigators providing
samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
for transgenic model; Xu et al., Nature Genetics 22, 37-44
(1999). Note: this is a NCI CGAP Library."
158 a 130 c 133 g 117 t

```

Query Match	39.48;	Score 94.6;	DB 11;	Length 538;
Best Local Similarity	62.48;	Pred. No. 4.5e-18;		
Matches 148;	Conservative	0;	Mismatches 89;	Indels 0;
			Gaps	0

OY 1 AAAATGCTGANTTCTGGTTGCTAGTACTCTCCAGATAATGAACCTCACTG 60
| | | | | | | | | | | | | | | |
Do 155 ACAAGCAGAGCTAATTAGTTGGCATCTCACCTGCACAATGAGAGCGACCCCTACACC 214
OY 61 AGCAGATGGCAGCTGTGTGAGAGAAGGCTGAAGGATGCCACTCACTCTGATTTTCATTG 120

Db	235	TTCA	GTAT	GGAC	AGTGT	GGGAC	ACAG	GAC	AGTAC	TAC	ACTT	CACT	CAAA	CTT	CT	CA	274									
QY	121	CAG	CAAAA	AGCT	TAC	CTGA	AT	TG	GAC	CA	CA	AG	TG	GG	CA	TTT	GC	180								
Eb	275	TCAC	TGAT	ACT	TGG	GTAT	CTAT	TG	AG	CCCC	CG	AG	GG	CA	GA	GTCT	TT	TC	AT	GA	TG	AG	TG	GG	CC	334

Cy	181	ATCTACGATGGGAGATTTTGACGAGTACAAATATGATGAGAAATTTCTACTTA	237
		..	
Db	335	ATCTCGGTGGGAGATTTTGATGATGATTAACCGGACCGCTTTTACACATTTCTA	391

RESULT 13

[illegible]

ORGANISM	Mus musculus
SOURCE	house mouse.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 674)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: rgsab@nci.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA sequencing by: Incyte Genomics, Inc. clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov
COMMENT	plate: LLM8785 row: e column: 10 high quality sequence stop: 607. Location/Qualifiers
FEATURES	1..674
SOURCE	/organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:3600753" /clone_lib="NCI-CGAP_Mam1" /tissue_type="tumor, biopsy sample" /dev_stage="3 months, virgin" /lab_host="DH10B" /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; Salt; Site:2; Note: Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT	198 a 166 c 165 g 145 t
ORIGIN	
Query Match	39.4%; Score 94.6; DB 10; Length 674;
Best Local Similarity	62.4%; Pred. No. 4.8e-18;
Matches 148; Conservative	0; Mismatches 89; Indels 0; Gaps 0;
CY	1 AAAATCCTCATTTTGTTGGTCTGAGTCACCTCCACAGTAATGATGAACCCCTACACTG 60
DB	253 ACAAAAGCAAGACTCATATTGTGGCGATCCCTCACCTGCACATGAGACGACCCTACACCC 311
CY	61 ACCAATGGGCAACTGTGGACGAGAAGGTGAAGAAGATCCACTCTCTGATTTCATG 120
DB	312 TTTCAGTATGACAGTAGTGGGACACAGGACATACATACACTTCACCTCCAACTCTCAG 371
CY	121 CAGGAAAAAATTTACTCTGATATGAGACCAACAAGTAGGAGCATTTGCCATGAGTGGGCTC 180
DB	372 TCACGATTAACCTTGGGTATCTATGTGAGCCCCGAGGAGAGACTCTTTGTCATGAGTGGGCC 431
CY	181 ATTCACGATGGGAGTATTTGACGAGTACAAATAATGATGAGAAAATTTCACTTATCCA 237
DB	432 ATCTCCGGTGGGAGATTTGATGTAGTAAACGTGAGCCGCGCTTCTTCATTTCTGA 488
RESULT 415	
LOCUS	BI251115 761 bp mRNA EST 17-JUL-2001
DEFINITION	602994311P1 NCI_CGAP_Mam5 Mus musculus CDNA clone IMAGE:5150186 5' /
ACCESSION	BI251115
VERSION	BI251115.1 GI:14800194
KEYWORDS	EST.
SOURCE	mouse;mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 761)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:36:15 ; Search time 521.98 Seconds
(without alignments)
104.132 Million cell updates/sec

Title: US-09-049-696-3

Perfect score: 240
Sequence: 1 AAATGCTGATGTTGTTGTT.....GAATCTACTATCAATG 240

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Dackfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	109.8	45.8	401 4 US-09-221-298-34	Sequence 34, Appl
2	34	5319	1 US-08-169-927-1	Sequence 1, Appl
3	32.4	13.5	1423 4 US-08-916-576B-3	Sequence 3, Appl
4	30.6	12.8	4403765 4 US-09-103-840A-2	Sequence 2, Appl
5	30.6	12.8	4411529 4 US-09-103-840A-1	Sequence 1, Appl
6	30	12.5	302 4 US-08-916-576B-19	Sequence 19, Appl
7	30	12.5	396 2 US-08-465-380-12	Sequence 12, Appl
8	30	12.5	396 2 US-08-480-478-41	Sequence 41, Appl
9	30	12.5	396 2 US-08-486-397-12	Sequence 12, Appl
10	30	12.5	396 2 US-08-486-399-12	Sequence 12, Appl
11	30	12.5	396 2 US-08-461-965-12	Sequence 12, Appl
12	30	12.5	396 2 US-08-326-110A-41	Sequence 41, Appl
13	30	12.5	396 2 US-08-634-641-12	Sequence 12, Appl
14	30	12.5	396 3 US-09-249-471-12	Sequence 12, Appl
15	30	12.5	396 3 US-09-249-472-12	Sequence 12, Appl
16	30	12.5	396 3 US-09-249-451-12	Sequence 12, Appl
17	30	12.5	396 3 US-08-809-455-12	Sequence 12, Appl
18	30	12.5	396 3 US-09-249-461-12	Sequence 12, Appl
19	30	12.5	396 3 US-09-249-448-12	Sequence 12, Appl
20	30	12.5	4693 4 US-09-359-756-1	Sequence 1, Appl
21	28.8	12.0	771 4 US-08-991-789A-241	Sequence 241, App
22	28.6	11.9	1000 3 US-08-961-083-53	Sequence 53, Appl
23	28.4	11.8	1815 4 US-09-042-785A-24	Sequence 24, Appl
24	28.4	11.8	2186 3 US-08-959-382-1	Sequence 1, Appl
25	28.4	11.8	2612 4 US-09-042-785A-3	Sequence 3, Appl
26	28.4	11.8	2638 4 US-09-042-785A-22	Sequence 22, Appl
27	27.6	11.5	2071 1 US-08-393-985-15	Sequence 15, Appl

28	27.4	11.4	3607 1 US-08-647-351B-1	Sequence 1, Appl
29	27.2	11.3	4032 1 US-08-107-748-3	Sequence 3, Appl
30	27.2	11.3	4032 1 US-08-245-809-4	Sequence 4, Appl
31	27.2	11.3	4032 5 PCT-US92-01385-3	Sequence 3, Appl
32	26.8	11.2	3417 2 US-08-978-458-7	Sequence 7, Appl
33	26.8	11.2	3417 3 US-08-978-454-7	Sequence 7, Appl
34	26.8	11.2	3417 4 US-09-385-288-7	Sequence 7, Appl
35	26.8	11.2	3417 4 US-09-977-555-7	Sequence 7, Appl
36	26.8	11.2	87350 3 US-08-781-891-79	Sequence 79, Appl
37	26.6	11.1	2746 2 US-08-576-165-3	Sequence 3, Appl
38	26.4	11.0	7218 1 US-08-232-463-14	Sequence 14, Appl
39	26.2	10.9	1125 1 US-07-928-462-1	Sequence 1, Appl
40	26.2	10.9	1125 3 US-08-273-247-1	Sequence 1, Appl
41	26.2	10.9	1332 2 US-08-481-814A-3	Sequence 3, Appl
42	26.2	10.9	1489 3 US-08-836-582-1	Sequence 1, Appl
43	26.2	10.9	1489 3 US-09-265-566-1	Sequence 1, Appl
44	26	10.8	747 5 PCT-US91-06234A-6	Sequence 6, Appl
45	26	10.8	2217 4 US-09-244-314-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-221-298-34
Sequence 34, Application US/09221298
Patent No. 6284241
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
FILE OF INVENTION: OF COLON CANCER
FILE REFERENCE: 210121.471
CURRENT APPLICATION NUMBER: US/09/221, 298
CURRENT FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 112
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 34
LENGTH: 401
TYPE: DNA
ORGANISM: Human
US-09-221-298-34

Query Match 45.8%; Score 109.8; DB 4; Length 401;
Best Local Similarity 92.9%; Pred. No. 1e-27;
Matches 170; Conservative 0; Mismatches 7; Indels 6; Gaps 5;

QY 1 AAATGCTGATGTTGTTGTTGC-TGACTTACTCTCCAGCTATGATGAACCTACACT 59
DB 215 aaatgctgattgttctggttgccttccacccagtgatgaaccttacct 274
QY 30 GAGCAGAT-GGGCAGCTGAGGAGGAGG--TGAAGATTCACCTCCTGATTTC 116
DB 275 gaggcagatgggcaacgtaggagaggaaggaggaagatccaccctcctgatttc 334
QY 147 ATTGCAAGAAAAAGTTAGC-TGAATATGACACCAAGGT-AGGGCATTGTCTAGT 174
DB 385 attgcagaaaagtttagcttgatgaatagaccacaaggaaggacttgcctgaat 394
QY 175 GGG 177
DB 395 ggg 397

RESULT 2
US-08-169-927-1
Sequence 1, Application US/08169927
Patent No. 5783441
GENERAL INFORMATION:
APPLICANT: Carl, Mitchell
APPLICANT: Dobson, Michael E.
APPLICANT: Ching, Wei Mei
APPLICANT: Dasch, Gregory A

TITLE OF INVENTION: Gene and Protein Applicable to the
 TITLE OF INVENTION: Preparation of Vaccines for *Rickettsia prowazekii* and
 TITLE OF INVENTION: *Rickettsia typhi* and the Detection of Both
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Counsel, Naval Medical R & D Command
 STREET: Bldg. 1, T-12, 8901 Wisconsin Ave.
 City: Bethesda
 STATE: MD
 COUNTRY: USA
 ZIP: 20889-5606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/169,927
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/742,128
 FILING DATE: 08/09/91
 ATTORNEY/AGENT INFORMATION:
 NAME: Spevack, A. David
 REGISTRATION NUMBER: 24,743
 REFERENCE/DOCKET NUMBER: 75,976
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 295-6759
 TELEFAX: (301) 295-1022
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5319 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: Internal
 ORIGINAL SOURCE:
 ORGANISM: *Rickettsia prowazekii*
 STRAIN: Breinl
 FEATURE:
 NAME/KEY: -35_signal
 LOCATION: 340..345
 FEATURE:
 NAME/KEY: -10_signal
 LOCATION: 363..368
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 391..5226
 FEATURE:
 NAME/KEY: RBS
 LOCATION: 379..386
 FEATURE:
 NAME/KEY: stem_loop
 LOCATION: 5270..5306
 PUBLICATION INFORMATION:
 AUTHORS: Carl, M.
 AUTHORS: Dobson, M. E.
 AUTHORS: Ching, W. M.
 AUTHORS: Daesch, G. A.
 TITLE: Characterization of the gene encoding the
 TITLE: protective S-layer protein of *Rickettsia*
 TITLE: *prowazekii*; presence of a truncated identical
 TITLE: homolog in *rickettsia typhi*
 JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
 DATE: 1990
 RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 5319

Query-Match	14.2%	Score 34	DB 1	Length 5319
Best local similarity	51.3%	Pred. No. 0.071		
Matches	79	Conservative	0	Mismatches 75; Indels 0; Gaps 0;
OY	3	AATGCTGATGTTCTGGTGGTCTAGTCTACTCTCCAGGTATATGATGAACCCCTACACTGANG	62	
Db	1810	AATATATATGCTCTCTCGACGCTGTCTTATTCAGTATAGGAGGAGTGCTAATATACCGGT	1869	
OY	63	CAGATGGGCAACTCTGGGAGAGAGAGGTGAAAGATCCACCTCCCTCGATTCTATTGCA	122	
Db	1870	GATATATGATGATACGGTGGTGGTTATATCTCTGGTTACAAACCACTTACTTATAGCTAACGATGCT	1929	
OY	123	GGAAAAAGTTAGCTGATATATGACACACAGGTA	156	
Db	1930	TCAAAAATTTATGACACTGATGCGGCCCAATATTA	1963	

```

1 RESULT .3
2 US-08-916-576B-3
3 : Sequence 3, Application US/08916576B
4 : Patent No. 6171816
5 :
6 : GENERAL INFORMATION:
7 : APPLICANT: YU, GUO-LIANG
8 : APPLICANT: DILLON, PATRICK J.
9 : APPLICANT: EENNER, REINHARD
10 : APPLICANT: ENDRESS, GREGORY A.
11 : TITLE OF INVENTION: NOVEL HUMAN GROWTH FACTORS
12 : NUMBER OF SEQUENCES: 45
13 : CORRESPONDENCE ADDRESS:
14 : ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
15 : STREET: 1100 NEW YORK AVENUE, SUITE 600
16 : CITY: WASHINGTON
17 :
18 : STATE: DC
19 : COUNTRY: US
20 : ZIP: 20005-3934
21 :
22 : COMPUTER READABLE FORM:
23 : MEDIUM TYPE: Floppy disk
24 :
25 : COMPUTER: IBM PC compatible
26 : OPERATING SYSTEM: PC-DOS/MS-DOS
27 : SOFTWARE: PatentIn Release #1.0, Version #1.30
28 : CURRENT APPLICATION DATA:
29 : APPLICATION NUMBER: US/08/916,576B
30 :
31 : FILING DATE:
32 : CLASSIFICATION: 536
33 :
34 : PRIOR APPLICATION DATA:
35 : APPLICATION NUMBER: US 60/024,347
36 : FILING DATE: 23-AUG-1996
37 :
38 : ATTORNEY/AGENT INFORMATION:
39 :
40 : NAME: STEFFE, ERIC K.
41 : REGISTRATION NUMBER: 36,688
42 : REFERENCE/DOCKET NUMBER: 1488.0500001
43 : TELECOMMUNICATION INFORMATION:
44 : TELEPHONE: (202) 371-2600
45 : TELEFAX: (202) 371-2540
46 :
47 : INFORMATION FOR SEQ ID NO: 3:
48 : SEQUENCE CHARACTERISTICS:
49 : LENGTH: 1423 base pairs
50 : TYPE: nucleic acid
51 : STRANDEDNESS: double
52 : TOPOLOGY: linear
53 :
54 : MOLECULE TYPE: CDNA
55 :
56 : FEATURE:
57 : NAME/KEY: CDS
58 : LOCATION: 88..603
59 :
60 : FEATURE:
61 : NAME/KEY: mat_peptide
62 : LOCATION: 157..603
63 :
64 : FEATURE:
65 : NAME/KEY: sig_peptide
66 : LOCATION: 88..156
67 :
68 : US-08-916-576B-3

```


RESULT 5
 US-09-103-840A-1
 : Sequence 1, Application US/09103840A
 : Patent No. 6294328
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: FLEISCHMAN, Robert D.
 : APPLICANT: WHITE, Owen R.
 : APPLICANT: FRASER, Claire M.
 : APPLICANT: VENTER, John C.
 : TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 : TITLE OF INVENTION: TUBERCULOSIS
 : FILE REFERENCE: 24366-20007.00

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,347
FILING DATE: 23-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0500001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA

Query_Match	12.5%; Score 30; DB 4; Length 302;
-------------	------------------------------------

QY 20 TCGTGAAGTCTACTCTCCAGGTAATGTAACCTTACCTGACAGATGGCAACTGTGG 79
DB 123 TGGAAGTCCGCTGGTATGATGACGATTAATACAGACAAATTTGGCAACTGTGT 182
QY 80 AGAGAAGGTGAAGAAGATCCACTCTGATTTTCATTGACAGAAAA 129
DB 183 TGAAGAAGCAATGCAACGATTAATGAGATTATTTACTTTTGACACAGAAA 232

RESULT 9

US-08-486-397-12
Sequence 12, Application US/08486397
Patent No. 586543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Ganssems, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 357
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Ancylostoma duodenale
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 10...237
US-08-486-397-12

Query Match 12.5%; Score 30; DB 2; Length 396;
Best Local Similarity 54.5%; Pred. No. 0.56;
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 20 TCGTGAAGTCTACTCTCCAGGTAATGTAACCTTACCTGACAGATGGCAACTGTGG 79
DB 123 TGGAAGTCCGCTGGTATGATGACGATTAATACAGACAAATTTGGCAACTGTGT 182

QY 80 AGAGAAGGTGAAGAAGATCCACTCTGATTTTCATTGACAGAAAA 129
DB 183 TGAAGAAGCAATGCAACGATTAATGAGATTATTTACTTTTGACACAGAAA 232

RESULT 10

US-08-486-399-12
Sequence 12, Application US/08486399
Patent No. 586543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Ganssems, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,399
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Ancylostoma duodenale
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 10...237
US-08-486-399-12

Query Match 12.5%; Score 30; DB 2; Length 396;
Best Local Similarity 54.5%; Pred. No. 0.56;
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 20 TCGTGAAGTCTACTCTCCAGGTAATGTAACCTTACCTGACAGATGGCAACTGTGG 79
DB 123 TGGAAGTCCGCTGGTATGATGACGATTAATACAGACAAATTTGGCAACTGTGT 182

QY 80 AGAGAAGGTGAAGAAGATCCACTCTGATTTTCATTGACAGAAAA 129
DB 183 TGAAGAAGCAATGCAACGATTAATGAGATTATTTACTTTTGACACAGAAA 232


```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Patient No. 05945275
GENERAL INFORMATION:
APPLICANT: GEORGE P. VLASUK; PATRICK ERIC APPLICANT: HUGO STANSSENS; JORIIS HILDA APPLICANT: LIEVEN MESSENS; MARC JOZEF APPLICANT: LAUREREYS; YVES RENE LAROCHE; APPLICANT: LAURENT STEPHANE JESPERS; and APPLICANT: YANNICK GEORGES JOZEFF

TITLE OF INVENTION: NEMATODE-EXTRACTED ANTI-TITLE OF INVENTION: COAGULANT PROTEIN
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,110A
FILING DATE: 18 OCTOBER 1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 208/290
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 base pairs
TYPE: nucleic acid
STRANDNESS: single
TOPOLOGY: linear

FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 10...237
OTHER INFORMATION:

US-08-326-110A-41

Query Match 12.5% Score 30; DB 2; Length 396;
Best Local Similarity 54.5%; Pred. No. 0.56;

Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0.

DQ 20 TGTCTGTCCTCCCTCCAGTAATGAACCACTGCAGTGGCATGTGG 79
 ||| | | | | | | | | | | | | | | | | | | | |
Db 123 TGAGAAGTCGCCCTTGCGATGCAATGACAGGATTTATCACAGAACAACTTTGGCAACTGTGT 182

DQ 80 AGAGAAGGATAAAGGATCCACCCTGCTCGTAGTTTCATTGCGAGAAAA 129
 ||| | | | | | | | | | | | | | | | | | | | |
Db 183 TCAGAAAAGCAATGCAACGATATGAGATATATCTTTGGACCAAGAAA 232

RESULT 13
US-08-634-641-12
Sequence 12, Application us/08634641
Patent No. 5955294
GENERAL INFORMATION:
APPLICANT: Vlasuk, George P. Vlasuk

[illegible]

```

1  NAME/KEY: Coding Sequence
2  LOCATION: 10...237
3  LENGTH: 366 base pairs
4  TYPE: nucleic acid
5  STRANDNESS: single
6  TOPOLOGY: linear
7  ORIGINAL SOURCE:
8  ORGANISM: Ancylostoma duodenale
9  FEATURES:
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
102
```


PI	Holroyd KJ, Levitt RC, Maloy WL, Lounched J, McLane M;
PI	Nicolaides NC, Zhou Y, Dong Q;
DR	WPI: 1999-550979/46.
DR	P-PSDB; AAB/4824.
XX	
PT	New nucleic acid encoding calcium activated chloride channel, used to
PT	identify, e.g. specific modulators for treating atopic allergy -
XX	
PS	Claim 1; Fig 4B; 75pp; English.
XX	
CC	The present sequence encodes the human interleukin 9 (IL-9) induced
CC	calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
CC	have anti-allergic, anti-asthmatic, anti-inflammatory and
CC	immunomodulatory activities. Compounds (A) that downregulate ICACC are
CC	used to alleviate asthma (or more generally atopic allergy), while those
CC	(B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
CC	inflammatory bowel disease (IBD) (or other autoimmune diseases).
CC	Measuring levels of ICACC-1 can be used in the diagnosis of asthma
CC	(increased levels) or IBD (reduced levels), also for monitoring
CC	treatment of these conditions. The ICACC proteins can be used:
CC	(i) to raise specific antibodies (Ab), useful: (a) as immunoassay
CC	reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
CC	(A) to disrupt bonding between ICACC and its ligands; and (iii) to
CC	identify modulators and binding partners. ICACC polynucleotides can be
CC	used to generate transgenic animals or recombinant cells, used to screen
CC	for antagonists, also as a source of therapeutic antisense agents or
CC	diagnostic probes (for quantifying mRNA expression, e.g. for
CC	identification of modulators).
XX	
SO	Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;
Query Match 100.0%; Score 240; DB 20; Length 2745;	
Best Local Similarity 100.0%; Pred. No. 2.4e-69;	
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AAAATGCTGATTTTGTGTTGCTGAGTCACTCCGACAGTAAATGAAACCTACACNG 60
DB	
QY	299 aaaaagctatgtctcgtgtcgtgctgagctcctccagtaatgaaacctacacg 358
DB	
QY	61 AGCAGATGGGCAACTGTGTGAGAGAAAGGCGTGAAGATCCACCTCACTCTGATTCATNG 120
DB	
QY	121 CAGGAAAAAGTTAGCTGAATATGACACCAAGATAGGCGCATTTGTCCATGAGTGGCTC 180
DB	
QY	419 caggaaaaaagttagctgtaataatgacccaagaagtaggcatctgtccatgagtggtc 478
DB	
QY	181 ANCTACGATGGGAGATTTTGACGAGTACAAATTAATGATGAGAAATCTACTATCCATG 240
DB	
QY	479 actcaagatcgggagatcattcgaagatacaataatgataagaaattctcattatccaatg 538
XX	
RESULT 2	
AAH46124	
ID	AAH46124 standard; cDNA: 2825 BP.
XX	
AC	AAH46124;
XX	
DT	11-SEP-2001 (first entry)
XX	
DE	Human CLCA1 cDNA, SEQ ID NO:26.
XX	
KW	Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
KW	expression inhibition; antisense therapy; gene therapy;
KW	chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
SS	ss.
XX	
OS	Homo sapiens.
XX	
EH	Key
FT	CDS Location/Qualifiers
	23..2767

```

      /tag= "a
      /product= "Human CLCA1"
      /transl_except= (pos:476..478, aa:Lys)
XX
XX
PN WO200138530-A1.
PD
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000MO-JP08232.
PR 24-NOV-1999; 99JP-0333479.
PR 27-APR-2000; 2000JP-0127589.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
PI Nakanishi A, Morita S;
DR WPI; 2001-355935/37.
DR P-PSTDB; AAB73716.
XX
PT New antisense nucleotide, useful for treatment and prevention of
PT bronchial asthma and chronic obstructive pulmonary disease -
XX
PS Example 5; Page 92-94; 104pp; Japanese.
XX
CC The invention relates to an antisense nucleotide targetted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents human CLCA1 cDNA.
S2
Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;

Query Match          100.0%; Score 240; DB 22; Length 2825;
Best local Similarity 100.0%; Pred. No. 2,4e-69;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 1 AAAATGCTGATGTTCGGTTCAGTAGCTACCTCACCAGTAGAATGATGAACCTACACTG 60
   |||||||
Db 321 aaaatcgtcgcttcggttcgagttccaccctcccaaggtaagaagaacctcaactg 380

OY 51 AGCAGATGGGCACACTGTGAGAGAAGGTGAAGAAGATCACCTCACTCGATTTCATTG 120
   |||||||
Db 381 agcagatgggcacactgygagagaaggygaagaagatcacctcactccltatttcagt 440

OY 121 CAGAAAAAAGTTAGTCGATATGAGACCCACAAGGTAGAGGCAATTTGCCATGAGTGGGCTC 180
   |||||||
Db 411 cagaaaaaaagttagctggaatalagaccacaagglagggcatcttgccaagtgggctc 500

OY 131 ATTCACGATGGGAGTATTTGACGAGTACATAATGATGAGAAATTCATTACTTATCATG 240
   |||||||
Db 501 attcagatggggagattttgcagtgagatacataatgatgagaattctactatccaatg 560

RESULT 3
AAH34879.
ID AA34879 standard; CDNA; 2854 BP.
AC
XX AAH34879;
XX
LI 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KV colorectal carcinoma; chromosome 1; ss.
XX
```


OS	Homo sapiens.
XX	
PN	W0200122920-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000WO-US26524.
XX	
PR	29-SEP-1999; 99US-0157137.
XX	
PA	03-NOV-1999; 99US-0163280.
XX	
PI	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Barash SC, Birse CE, Rosen CA;
XX	
DR	WPI: 2001-235357/24.
XX	
DR	P-PSDB; AAG75474.
XX	
PT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX	
PS	useful for preventing, diagnosing and/or treating colorectal cancers -
XX	
PS	Claim 1; Page 3462-3463; 9803pp; English.
XX	
CC	AAB32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX	
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where
XX	
CC	the proteins are collectively known as colon cancer antigens. The colon
XX	
CC	cancer antigens have cytostatic activity and can be used in gene
XX	
CC	therapy and vaccine production. N and P may be used in the prevention,
XX	
CC	diagnosis and treatment of diseases associated with inappropriate P
XX	
CC	expression. For example, N and P may be used to treat disorders
XX	
CC	associated with decreased expression by rectifying mutations or deletions
XX	
CC	in a patient's genome that affect the activity of P by expressing
XX	
CC	inactive proteins or to supplement the patients own production of P.
XX	
CC	Additionally, N may be used to produce the colon cancer-associated ps,
XX	
CC	by inserting the nucleic acids into a host cell and culturing the cell
XX	
CC	to express the proteins. N and P can be used in the prevention, diagnosis
XX	
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX	
CC	and AAB77789 represent sequences used in the exemplification of the
XX	
CC	present invention.
XX	
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX	
CC	missing at time of publication, meaning no sequences are present for
XX	
CC	SEQ ID NO:1027 to 1052, 7921 and 7922.
XX	
SO	Sequence 2854 BP: 885 A; 628 C; 643 G; 698 T; 0 other:
XX	
Query Match	100.0%; Score 240; DB 22; Length 2854;
Best Local Similarity	100.0%; Pred. No. 2,4e-69;
Matches 240; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1 AAAATGCTGANGTTTGTGTTGCTGAGTCACCTCAGATAATGATGAACCCCTACACTG 60
Db	333 aaaagctgcatgtcttgcgtgtcgtgctacccctccagtaagtatgaacctactac 392
OY	61 AGCAATATGGGCACACGTGGAGAGAGGGGAAAGATGCCACCATCTCGATTTCATG 120
Db	333 agcagaatggcaaccgtggagaaagggtgaaggatccaaccctcctgatattcatty 452
OY	121 CAGGAAAAAAGTTTAGCTGAATATGAGCCACAAGGCTGGGGCATTTGTCCATGAGTGGGCTC 180
Db	453 caggaaaaaagttagctgaataatagcaccaacaagtgtagggcatttgcattgagtggttc 512
OY	181 ATCTACGATGGGGAGTATTTGACGAGTACAAATTAATATGAGAATTTCTACTTATCCAATG 240
Db	513 atctacgatgggagatatttgacgagttacataaatatgagaaattctactatcataatg 572
RESULT 4	
ID	AAF81787
XX	
KC	AAF81787 standard; CDNA: 2854 BP.
XX	
XX	AAF81787;
XX	

12: UN-2001 (first entry)
Human secreted protein gene 1 SEQ ID NO:11.
Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
dermatological; immunosuppressive; anti-HIV;
immunostimulant; cyostatic; cardiant; vascular; anti-angiogenic;
ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
antialzheimer; antiparkinsonian; antimicrobial; vulnary; gene therapy;
immune disorder; hyperproliferative disorder; cardiovascular disease;
cancer; angiogenic disorder; neurological disorder; infectious disease;
wound healing; regeneration; chemotaxis; chromosome 1; ss.
Homo sapiens;
WO300112775-A2.
22-FEB-2001.
16-AUG-2000; 2000MO-US22325.
17-AUG-1999; 9905-0149182.
(HUMA-) HUMAN GENOME SCI INC.
Roisen CA, Ni J, Florence KA, Fiscella M, Wel P, Baker KP;
Bisce CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;
WPI; 2001-147550/15.
P-PSDB; AAB74733.
Nucleic acids encoding 25 human secreted polypeptides, useful for
preventing, diagnosing and/or treating e.g. cancers, Parkinson's
disease and diabetic retinopathy -
Claim 1; Page 441; 485pp; English.
AAB71787 to AAB71817 encode the human secreted proteins given in AAB74733
to AAB74772. Human secreted proteins can have activities based on the
tissues and cells they are expressed in. Example of activities include:
immunomodulatory; antisclerotic; dermatological; immunosuppressive;
anticonvulsant; anti-HIV; immunostimulant; cyostatic; cardiant;
vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and
vulnerary. Human secreted proteins can be used in gene therapy and
vaccines. Human secreted protein nucleotide sequences (NAMI) and proteins
(PBI) may be used in the prevention, diagnosis and treatment of diseases
associated with inappropriate polypeptide expression. For example, NAMI
and PBI may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patients genome
that affect the activity of proteins by expressing inactive proteins or
to supplement the patients own production of polypeptides. Disorders that
may be prevented, diagnosed and/or treated include immune disorders,
hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
angiogenic disorders, neurological disorders, infectious diseases and/or
for promoting wound healing, regeneration and/or chemotaxis. AAB7178 to
AAB71786 and AAB74732 represent sequences used in the exemplification of
the present invention.
Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;
Query: Match 100.0%; Score 240; DB 22; Length 2854;
Best Local Similarity 100.0%; Pred. No. 2,4e-69;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0
11 AAAAGCCATGTCGTTGGTGGTGGAGCTACCTCCAGGATATGATGAAACCTAACG 60
333 aaaaagcgtatgtctgtgtctgtgtcgtacactccctccaggtatgatgaaccctaccc 352
151 ACCAGATGGGCACTCTGAGAGAAAGGTAAGATCCACCTCACTCCTGATTTCAATG 120
353 accaagatgggcaactctgtgagagaaaggtaaagagatccacccctcctgatttcattg 452
353 accaagatgggcaactctgtgagagaaaggtaaagagatccacccctcctgatttcattg 452

XX	27-APR-2000; 2000Jp-0127589.
PA	(TAKE) TAKEDA CHEM IND LTD.
XX	
PI	Nakanishi A, Morita S;
XX	
DR	MPI: 2001-355935/37.
XX	P-PSTD: AAB7715.
PT	New antisense nucleotide, useful for treatment and prevention of
PT	bronchial asthma and chronic obstructive pulmonary disease -
XX	
PS	Claim 3; Page 80-82; 104pp: Japanese.
XX	
CC	The invention relates to an antisense nucleotide targeted to the mouse
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC	the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC	relates to an antibody specific for the Gob-5 protein, medical and
CC	diagnostic compositions containing the antisense nucleotide or the
CC	antibody, and methods and kits for screening for compounds which inhibit
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC	The antisense oligonucleotides and antibody are therefore useful for the
CC	treatment and prevention of bronchial asthma and chronic obstructive
CC	pulmonary disease. The present sequence represents the mouse Gob-5
CC	gene coding sequence.
XX	
SQ	Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other:
Query Match	62.0%; Score 148.8; DB 22; Length 2739;
Best Local Similarity	76.2%; Pred. No. 3.3e-39;
Matches 183; Conservative	0; Mismatches 57; Indels 0; Gaps 0;
YY	1 AAAATGCTGAAGTTCGTGGTGCTGACTACTCCTCCAGGTAAATGATGAACCTCACACTG 60
Dbb	
YY	299 aaaagcgcgtagtctctgtatcaaaacacgacctcttgacgatgagccctaacg 358
YY	61 AGCAGATGGCACTGTGCGAGAGAAGGTTGAAGSATTCCACCCTCACCTCATTTTCATTG 120
Dbb	
YY	359 aacatatagagcatgtggaagaaggagatcaagattcaactgacctgaccttcttag 418
YY	121 CAGGAAAAAGCTTAGCTGAATATGAGACAACAAGTGAAGGCATTGTGCCATGAGTGGGCTC 180
Dbb	
YY	419 caggaaaagaagcgtgactgacagtgggccacaagacagagaccttgcctcgagtggttc 478
Dbb	
YY	181 ATCTACGATGGGAGTATTGTGACGAGTACCAATAATGATGAGAAATTCTACTATCCAAATG 240
Dbb	
YY	479 acttcgcgtggggagtgcttaatagtatacaaaagaagagagtgcttactatccaag 538
Dbb	
RESULT 11	
AAH46120	
ID	AAH46120 standard; cDNA; 2843 BP.
XX	
XX	AAH46120;
DT	11-SEP-2001 (first entry)
XX	
DE	Mouse Gob-5 cDNA, SEQ ID NO:22.
XX	
KW	Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
KW	expression inhibition; antisense therapy; gene therapy; bronchial asthma;
XX	chronic obstructive pulmonary disease; antiasthmatic; ss.
OS	
XX	Mus sp.
FH	
FT	key Location/Qualifiers
CDS	15..2756
FT	/*tag= a
XX	/product= "Mouse Gob-5"
NN	
XX	WO200138530-A1.

```

XX 31 MAY -2001 .
XX 22 NOV -2000; 2000MO-JP08232.
XX 24 NOV -1999; 99JP-0333479.
XX 27 APR -2000; 2000JP-0127589.
XX (JUNKE ) TAKEDA CHEM IND LTD.
XX NAKANISHI A, MORITA S;
XX MFI; 2001-355935/37.
XX B-PDSB; AAB73715.
XX New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease
XX Example 1; Page 89-91; 104pp; Japanese.
XX The invention relates to an antisense nucleotide targetted to the mouse
XX Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX relates to an antibody specific for the Gob-5 protein, medical and
XX diagnostic compositions containing the antisense nucleotide or the
XX antibody, and methods and kits for screening for compounds which inhibit
XX the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX The antisense oligonucleotides and antibody are therefore useful for the
XX treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
XX Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other:
SD
Query Match 62.0%; Score 148.8; DB 22; Length 2843;
Best Local Similarity 76.2%; Pred. No. 3,4e-39;
Matches 183; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY 1 AAAATGCGAGTGTTCGGTTGCTGCTACTCCCTCCAGGTAAATATGTAACCCCTACACTG 60
Eb 1 ||| ||||| || | | | ||| ||| ||||| ||||| |
Dy 333 aaagcgatgagtgccttgatcaacaacaccgccctcagaactgaatgagccctaacccg 372
QY 51 AGCAGATGGGCACAACGTGGAGAGAAGGGTGAAAAGCATGCCACTCCTGATTTCATTG 120
Eb 1 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |
Dy 273 aacataatggagcatgttgygaaaagggaalcatcagattccactgcaccttccttag 432
QY 121 CAGGAATAAAGTTAGCTGAATATGGACCAACAAGGTAGGGCATTTGTCCATGAGTGGCTC 180
Eb 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |
Dy 433 caggaagaagcatgctcagtaatggccccaaagaacagacctttcatcagtgtggtc 492
QY 111 ATCTACGATGGGAGATATTTCAGCAGTCAATTAATGATGAGCAAAATTCTACTTACCATG 240
Eb 1 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |
Dy 493 acttcgatgggagtgctttaatgaatacacacaacagcagagaagttctaataccaag 552
RESULT : 12
ID AF81925 standard; CDNA: 2931 BP.
AF81925;
MA81925;
LT 13 JUN -2001 (first entry)
DE Mutine ICACC-1 nucleotide sequence.
XX ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
XX interleukin-9 induced calcium activated chloride channel; IL-9;
XX calcium activated chloride channel; anti-allergic; anti-asthmatic;
XX anti-inflammatory; immunomodulatory; cystic fibrosis;
XX inflammatory bowel disease; autoimmune disease; ss.
XX Must sp.
XX Location/Qualifiers
```



```
FT CDS 8..2749
FT FT /*tag= a "ICACC-1"
FT FT /note= "IL-9 induced calcium activated chloride channel"
XX MO9944620-A1.
XX PD 10-SEP-1999.
XX PF 03-MAR-1999; 99WO-US04703.
XX PR 03-MAR-1998; 98US-0076815.
XX PA (MAGA-) MAGALININ PHARM INC.
XX PI Holroyd KJ, Levitt RC, Maloy WT, Lounahed J, McLane M;
XX PI Nicolaidis NC, Zhou Y, Dong Q;
XX DR WPI; 1999-550979/46.
XX DR P-PSDB; AAB74822.
XX PT New nucleic acid encoding calcium activated chloride channel, used to
XX PT identify, e.g. specific modulators for treating atopic allergy -
XX PS Claim 2; Fig 2; 75pp; English.
XX CC The present sequence encodes the murine interleukin 9 (IL-9) induced
XX CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
XX CC have anti-allergic, anti-asthmatic, anti-inflammatory and
XX CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
XX CC used to alleviate asthma (or more generally atopic allergy), while those
XX CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
XX CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
XX CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma
XX CC (increased levels) or IBD (reduced levels), also for monitoring
XX CC treatment of these conditions. The ICACC proteins can be used:
XX CC (1) to raise specific antibodies (Ab), useful: (a) as immunoassay
XX CC reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as
XX CC (A) to disrupt bonding between ICACC and its ligands; and (iii) to
XX CC identify modulators and binding partners. ICACC polynucleotides can be
XX CC used to generate transgenic animals or recombinant cells, used to screen
XX CC for antagonists, also as a source of therapeutic antisense agents or
XX CC diagnostic probes (for quantifying mRNA expression, e.g. for
XX CC identification of modulators).
XX SQ Sequence 2931 BP; 867 A; 719 C; 684 G; 661 T; 0 other;

Query Match 61.3%; Score 147.2; DB 20; Length 2931;
Best Local Similarity 75.8%; Pred. No. 1.2e-38;
Matches 182; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

OY 1 AAAATGCGATGTTGGTGTCTGCTAGTACTCTCCAGTAATGATGAACCTACATG 60
DB 306 aaacgcgtagctcgtatcaacaaccagccctcgtgaaatgagagccctaacacg 365
OY 61 AGCAGATGGCACTGTGAGAGAGAGGCTGAAGGATCCACTCACTTATTTGATTG 120
DB 366 aacatagtagcgtatgtaggaaagggatcagatccactcgtcctcgtactctag 425
OY 121 CAGGAAAAAAGTTAGCTGAATATGAGCACAAGTAGGCGATTGTCCATGATGGCTC 180
DB 426 caggaaagaagcgtcgtatgtaggacacagaagaccttgcattcattagtgagctc 485
OY 181 ATTCAGATGGGAGTATTGACGAGTACATTAATGATGAGGAATTCACATTATCCATG 240
DB 486 actccgctggggagcttcaatgaatacaacaagcagagaagttcattcattcaaga 545

RESULT 13
AAZ65095
ID AAZ65095 standard; cDNA; 3265 BP.
XX
```

```
AC AAZ65095;
XX 05-APR-2000 (first entry)
XX DE Membrane-bound protein PRO1124 encoding cDNA.
XX KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
XX KW pharmacological; receptor immunoadhesin; gene mapping; ss.
XX OS Homo sapiens.
XX PN WO9963088-A2.
XX ED 09-DEC-1999.
XX PF 02-JUN-1999; 99WO-US12252.
XX PR 02-JUN-1998; 98US-0087607.
XX PR 02-JUN-1998; 98US-0087609.
XX PR 02-JUN-1998; 98US-0087759.
XX PR 03-JUN-1998; 98US-0087827.
XX PR 04-JUN-1998; 98US-0088021.
XX PR 04-JUN-1998; 98US-0088025.
XX PR 04-JUN-1998; 98US-0088028.
XX PR 04-JUN-1998; 98US-0088029.
XX PR 04-JUN-1998; 98US-0088030.
XX PR 04-JUN-1998; 98US-0088033.
XX PR 04-JUN-1998; 98US-0088326.
XX PR 05-JUN-1998; 98US-0088167.
XX PR 05-JUN-1998; 98US-0088202.
XX PR 05-JUN-1998; 98US-0088212.
XX PR 05-JUN-1998; 98US-0088217.
XX PR 05-JUN-1998; 98US-0088255.
XX PR 05-JUN-1998; 98US-0088722.
XX PR 10-JUN-1998; 98US-0088730.
XX PR 10-JUN-1998; 98US-0088734.
XX PR 10-JUN-1998; 98US-0088738.
XX PR 10-JUN-1998; 98US-0088740.
XX PR 10-JUN-1998; 98US-0088741.
XX PR 10-JUN-1998; 98US-0088742.
XX PR 10-JUN-1998; 98US-0088810.
XX PR 10-JUN-1998; 98US-0088811.
XX PR 10-JUN-1998; 98US-0088824.
XX PR 10-JUN-1998; 98US-0088825.
XX PR 10-JUN-1998; 98US-0088826.
XX PR 10-JUN-1998; 98US-0088828.
XX PR 11-JUN-1998; 98US-0088851.
XX PR 11-JUN-1998; 98US-0088853.
XX PR 11-JUN-1998; 98US-0088876.
XX PR 11-JUN-1998; 98US-0089090.
XX PR 12-JUN-1998; 98US-0089105.
XX PR 16-JUN-1998; 98US-0089440.
XX PR 16-JUN-1998; 98US-0089512.
XX PR 16-JUN-1998; 98US-0089514.
XX PR 17-JUN-1998; 98US-0089532.
XX PR 17-JUN-1998; 98US-0089538.
XX PR 17-JUN-1998; 98US-0089598.
XX PR 17-JUN-1998; 98US-0089599.
XX PR 17-JUN-1998; 98US-0089600.
XX PR 17-JUN-1998; 98US-0089653.
XX PR 18-JUN-1998; 98US-0089801.
XX PR 18-JUN-1998; 98US-0089907.
XX PR 18-JUN-1998; 98US-0089908.
XX PR 19-JUN-1998; 98US-0089947.
XX PR 19-JUN-1998; 98US-0089948.
XX PR 19-JUN-1998; 98US-0089952.
XX PR 22-JUN-1998; 98US-0090246.
XX PR 22-JUN-1998; 98US-0090252.
XX PR 22-JUN-1998; 98US-0090254.
XX PR 23-JUN-1998; 98US-0090349.
XX PR 23-JUN-1998; 98US-0090355.
XX PR 24-JUN-1998; 98US-0090429.
XX PR 24-JUN-1998; 98US-0090431.
```


XX Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;

Query Match 54.5%; Score 130.8; DB 22; Length 3265;
 Best Local Similarity 73.0%; Pred. No. 3.2e-33;
 Matches 168; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```

OY 1 AAAATGCTGATGTTGCTGCTGAGTCTACTCTCCAGTATGATGAACCTTACACTG 60
    ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 320 aacatgctgagttagttagtgaccacactccacgtagagatgacacataacca 379
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 61 AGCAGATGGGCACTGTGAGAGAGAGGTGAAGATCCACTCCTCTGATTTCATTG 120
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 380 agcagtlcacgaatgtagagaaagcgaaatacattcaaccctgacctctac 439
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 121 CAGGAAAAAGTTAGTGAATATGGCCACCAAGTAGGSCATTTGCCATGAGTGGGCTC 180
    ||||||| | ||||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 440 ttggaataaataaataatgaaatgacacacagcaactgttgcacatgagtggttc 499
    ||||| | ||||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 181 ATCTAGATGGGAGTATTGTGACGATACAAATATGATGAGAAATTCCTAC 230
    ||||| | ||||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 accctcggtgggaggtgtttagtgatgaatgaagaatcagccttctac 549
    ||||| | ||||||| ||||| ||||| ||||| ||||| ||||| |||||
    
```

Search completed: April 4, 2002, 03:17:19
 Job time: 60191 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 02:56:19 ; Search time 22700.8 Seconds

(without alignments)
174.413 Million cell updates/sec

Title: US-09-049-696-3

Perfect score: 240
Sequence: 1 AAATGCTGATGTTCTGCTT.....GAAATTCCTACTATCCAAATG 240

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenBankl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_lm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_lm: *
20: em_ov: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vi: *
30: em_hcgo_hum: *
31: em_hcgo_inv: *
32: em_hcgo_rnd: *
33: em_hcgo_hum: *
34: em_hcgo_inv: *
35: em_hcgo_rnd: *
36: em_hcgo_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	240	100.0	2826	9	AF127036	AF127036 Homo sapi
2	238.4	99.3	3311	9	AX193489	AX193489 Sequence
3	238.4	99.3	3311	9	AF039400	AF039400 Homo sapi
4	169.2	70.5	3079	9	AF095584	AF095584 Sus scrofa
5	153	63.8	35278	9	AF039401	AF039401 Homo sapi
6	153	63.8	113764	9	HS1651E10	AL122002 Human DNA
7	153	63.8	164891	2	AL358950	AL358950 Homo sapi
8	148.8	62.0	2937	10	AB017156	AB017156 Mus muscu
9	130.8	54.5	3204	9	AF127035	AF127035 Homo sapi
10	130.8	54.5	3221	9	AK000072	AK000072 Homo sapi
11	130.8	54.5	3265	6	AX092338	AX092338 Sequence
12	129.8	45.8	401	6	AX192467	AX192467 Sequence
13	99.6	41.5	3415	9	AF043976	AF043976 Homo sapi
14	98	40.8	2984	4	BT036445	BT036445 Bos taurus
15	94.8	39.5	1265	4	AF001264	AF001264 Bos taurus
16	94.8	39.5	2820	4	AF001263	AF001263 Bos taurus
17	94.8	39.5	3288	4	AF001262	AF001262 Bos taurus
18	94.8	39.5	3317	4	AF001261	AF001261 Bos taurus
19	94.6	39.4	3022	10	AF15852	AF15852 Mus muscu
20	94.6	39.4	3058	10	AF047838	AF047838 Mus muscu
21	94.6	39.4	3137	10	AF108501	AF108501 Mus muscu
22	94.6	39.4	3471	10	BC008147	BC008147 Mus muscu
23	94.6	39.4	3471	10	AF052746	AF052746 Mus muscu
24	77.6	32.3	2832	9	AF127980	AF127980 Homo sapi
25	77.6	32.3	2970	9	AF043977	AF043977 Homo sapi
26	77.6	32.3	4077	6	AX054697	AX054697 Sequence
27	76	31.7	3604	9	AB026833	AB026833 Homo sapi
28	59	24.6	140718	2	AL356270	AL356270 Homo sapi
29	59	24.6	175591	2	AC068071	AC068071 Homo sapi
30	52	21.7	140718	2	AL356270	AL356270 Homo sapi
31	36	15.0	19436	1	U67522	U67522 Methanococ
32	35.6	14.8	142012	9	AL591428	AL591428 Human DNA
33	34.8	14.5	124096	2	AL354988	AL354988 Homo sapi
34	34.8	14.5	156069	9	AC079256	AC079256 Homo sapi
35	34.8	14.5	178714	33	AC022925	AC022925 Homo sapi
36	34.8	14.5	182304	2	AC092625	AC092625 Homo sapi
37	34.6	14.4	1251	4	MEBGK1	X64296 M. eugeni
38	34.6	14.4	2203	14	REO30CPMDA	REO30CPMDA
39	34.6	14.4	2203	14	RSU24260	RSU24260 Reovirus ty
40	34.6	14.4	2207	14	REO3M2C	REO3M2C Reovirus se
41	34.6	14.4	4851	1	AF123718	AF123718 Rickettsi
42	34	14.2	5015	1	AF161079	AF161079 Rickettsi
43	34	14.2	5015	1	AF211820	AF211820 Rickettsi
44	34	14.2	5015	1	AF211821	AF211821 Rickettsi
45	34	14.2	5131	1	R1RSPAP	M37647 Rickettsia

ALIGNMENTS

RESULT 1
AF127036
LOCUS Homo sapiens calcium-activated chloride channel protein 1 (CACCL)
DEFINITION mRNA, complete cds.
ACCESSION AF127036
VERSION AF127036.1
KEYWORDS GI:4585468
SOURCE human
ORGANISM Homo sapiens

REFERENCE
AUTHORS Agnelli, M., Verma, T., and Culouscou, J. M.
TITLE Identification of three novel members of the calcium-dependent chloride channel (CACCL) family predominantly expressed in the digestive tract and trachea

JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PubMed 10437792

FEATURES	source	Location/Qualifiers
BASE COUNT	1028 a 692 c 742 g 849 t	
ORIGIN		
Query Match	99.3% Score 238.4; DB 9; Length 3311;	
Best Local Similarity	99.6%; Pred. No. 1.7e-64;	
Matches 239; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
Oy	1 AAAAGCTGATGTTGTTGTTGCTGCTGCTACCTCCCGAGTAAATGATGAAACCTPACACG 60	
Db	650 AAAATGCTGATGTTGTTGTTGCTGCTGCTACCTCCCGATATGATGAAACCTTACACG 709	
Oy	61 AGCAGATGGCGCAACGTGTGAGAGAAAGGGTGAAGAAGATCCACCTCACTCCTGATTTTCATTG 120	
Db	710 AGCAGATGGCGCAACGTGTGAGAGAAAGGGTGAAGAAGATCCACCTCACTCCTGATTTTCATTG 769	
Oy	121 CAGGAAAAAAGTTAGCTGAATATGACACCAAGTAAGGCAATTTTCATGATGATGGGCTC 180	
Db	770 CAGGAAAAAAGTTAGCTGAATATGACACCAAGTAAGGCAATTTTCATGATGATGGGCTC 829	
Oy	181 ATCTCAGATGGGGAGTATTTTGACGCGTCAATAATGATGAGAAATTCATATTCGAATG 240	
Db	830 ATCTCAGATGGGGAGTATTTTGACGCGTCAATAATGATGAGAAATTCATATTCGAATG 889	
RESULT 4		
AF095584	3079 bp mRNA	MAM 13-OCT-2000
LOCUS	Sus scrofa epithelial chloride channel protein (ABCC) mRNA,	
DEFINITION	complete cds.	
ACCESSION	AF095584	
VERSION	AF095584.1	
KEYWORDS	GI:6002645	
SOURCE		
ORGANISM	Sus scrofa	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.	
TITLE	1 (bases 1 to 3079)	
JOURNAL	Cloning a chloride conductance mediator from the apical membrane of	
MEDLINE	porcine ileal enterocytes	
PubMed	Physiol. Genomics (Online) 3 (2), 101-111 (2000)	
REFERENCE	11015605	
AUTHORS	2 (bases 1 to 3079)	
TITLE	Gapar,K.J., Racette,K.J., Gordon,J.R., Loewen,M.E. and	
JOURNAL	Forsyth,G.W.	
REFERENCE	Cloning a chloride conductance mediator from the apical membrane of	
AUTHORS	porcine ileal enterocytes	
TITLE	Physiol. Genomics (Online) 3 (2), 101-111 (2000)	
JOURNAL	Submitted (28-SEP-1998) Veterinary Physiological Sciences,	
REFERENCE	University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4,	
AUTHORS	Canada	
FEATURES	Location/Qualifiers	
source	1..3079	

/organism="Sus scrofa"
 /db_xref="taxon:9823"
 /tissue_type="ileal mucosa"
 1. .3079
 /gene="AECc"
 133. .2886
 /gene="AECc"
 /function="Involved in exocrine chloride channel conductance"
 /codon_start=1
 /product="epithelial chloride channel protein"
 /protein_id="AAP00077.1"
 /db_xref="GI:6002646"
 /translation="MGSRFSLFTIVHLLGASNSLIQUNGSGYEGIVADPNP
 EDELKIKDMVTKASFLYLEAEAEKRYEFNVNLIIPASKMAPEYKPLEYKNA
 DVVTEPEPPENDPYEIQMGNGCEKEKLEIFTEDPAAKVKVLDGQRYFHEMAH
 LRMCVFNENNEBCEFLYSNKKRQEVICSAALRGYNVLPCOGGSCVTRPCRADVTGIL
 FORKECEFLPDEOGEKASIMPAOSIDIVVEFCRKNHKEPNQONCKLRSTVEVY
 QDSDFKTPPTPTPPAPPEPSILQIRYICLVLDGSGMTVGGRKRLNOACKLEFL
 LQYDEGAMVMVAFDSAAVYKSLVQINSAEEDLARSILPTAASGTSVCSLRNS
 FTVKIKKIPDSEVLITLDGEDNTISCFEYVQNGALHTVALGSAAELELSO
 MTGILQYASDAQANNGLIDAFGALSSGNRAASORSIQLESGLTLQNNEMNCTVVA
 DYSQKOTFLPLITERKFLSPIPEFVPSGSDSFLGKHNKMAAYRPGCTAAGVM
 KYSIOASSQLETLTVSSRRSSATLPPTVYSSKMDKSGFPPSPVYTKIHQGLPIIL
 RAKYATLISENGKTVLEILDNGAGADATNDGDISYFPAAYANRYSIKVWALGG
 VNTRRRARPPMSAMIRKVIENGELKMPNPRDINKDLOGQVCFSPRIASGSPY
 ASDVPKSTIPDLFPCKKITDLKAGIQGNLNLWIAAGDDYDHRADRTITRISTN
 LDLDKFNDSVQVMTDLIPKEANSEEFYVFKPGCIPPTNCTDLFLAVQADVKNILMS
 EISNIAQVSLFLPEAPEPEPEPEPPAPSLPCPEIOVNSTIPGIIHLKIMKWLGEIOL
 SIA"

BASE COUNT 862 a 758 c 755 g 674 t
 ORIGIN

Query Match 70.5%; Score 169.2; DB 4; Length 3079;
 Best Local Similarity 81.9%; Pred. No. 1.2e-42;
 Matches 195; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 AAAATGCTGATGTTGTTGCTGCTGAGTCTACTCTCCAGATATGATGAAACCTTCACTG 60
 Db 411 AAAATGCTGATGTTGTTGCTGCTGAGTCTACTCTCCAGATATGATGAAACCTTCACTG 490
 QY 61 AGCAGATGGGCAACGTGTGGAGAGGAAGGGTGAAGAGATCCACTCCTCTGATTTCATTG 120
 Db 411 AGCAGATGGGCAACGTGTGGAGAGAGGAAGGGTGAAGAGATCCACTCCTCTGATTTCATTG 550
 QY 111 CAGGAAAAAAGTTACCTGAATATGAGCACCACAAGTAGGAGCATTTGTCATGAGTGGGCTC 180
 Db 511 CAGGAAAAAAGTTACCTGAATATGAGCACCACAAGTAGGAGCATTTGTCATGAGTGGGCTC 610
 QY 111 ATCTACGATGGGGATATTGTCAGAGTACATATGATGAGAAATTCATTATCCAA 238
 Db 611 ACCTCGCATGGGGCGCTGTTTATGAGTACACACAGCAGACAGCAAGTTCATTATCCAA 668
 RESULT 5
 AF039401 AF039401
 LOCUS 1 AF039401 35278 bp DNA PRI 14-DEC-1998
 DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hclcal) gene,
 complete cds.
 ACCESSION AF039401
 VERSION 1
 KEYWORDS AF039401.1 GI:4009459
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 35278)
 Gruber,A.D., Elble,R.C., Ji,H.L., Schreuer,K.D., Fuller,C.M. and
 Pauli,B.U.
 Genomic cloning, molecular characterization, and functional
 analysis of human CLCA1, the first human member of the family of
 Ca2+-activated Cl- channel proteins
 Genomics 54 (2), 200-214 (1996)


```

MEDLINE 99047526
REFERENCE 2 (bases 1 to 35278)
AUTHORS Gruber,A.D., Eibler,R. and Pauli,B.U.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1997) Department of Pathology, College of
Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES
Source
1.35278
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
1.1617
/gene="hclcal"
1.33522
/gene="hclcal"
767.1265
/rpl_family="L1 transposable element"
/rpl_type="dispersed"
1587.1596
/gene="hclcal"
/note="putative"
1618.1708
/gene="hclcal"
/number=1
join(1618..1708,1962..2383,6667..6807,6979..7126,9696..9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33522)
/gene="hclcal"
/product="calcium-dependent chloride channel-1"
1962..2383
/gene="hclcal"
/number=2
join(2222..2383,6667..6807,6979..7126,9696..9801,15454..15631,18591..18809,19783..20010,22253..22427,24522..24628,26644..26859,27447..27708,28766..28936,31840..32079,32919..33522)
/gene="hclcal"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95429.1"
/db_xref="GI:4009460"
/translation="MGPFKSSVPILILHLEGALNSLIQLNNNGEGLVVAIDPNV
EDETLIQIKDMTQASLYLEATGKRFYKNAIILPETWTKADYVRKLETKNA
DVLVAESTPGNDPEPTQMGKCEGKGRHILPDTAGKLAEGYGGKAFYHENA
LRMGVDEYNNDEKFTLSNGRIQAVRCSAGITGTNYKKCGGSCYTRCTENKVTGL
YKGCPEVLQSRTEKASIMFAQVDSIVFCEBONHNKRAPKONKCNLRSTWYEL
ROSEDFKKTTPMTQPPNPTFSLQIGQRLVCLVDKSGSMATGNRLNRNLQAGFL
LQTVELGSWGVMTFDSAHOSELQIINGSDRDLAKRLPAASGSAIECEGLRSA
FTVIRKRYPTDGEIYLLDGEDNTISGCNEVKOSGAILHTPAAGSAOEMLELSK
MTGGLQTVASDOYNNGLIDAFGALSNGNAVSQLESKLTLLQNSAOEMLETVY
DSVVGDTLELITWTOPOQIILMDPSGGOGFEVYDKNKMYLQIPGAKGTWY
SIQASSOTLITVTSRASNATLIPITVTSNTNDSKPSPLVYANIRGASPIILRA
SVTALIESVNGKTVTLNDNGAGADKTDGVSRFTYTDINGRSVAVRALGVAN
AARRVLPQOSGALYIPGLTENDEIQMNPRPINDDVOHKVCFSSRSGSPVAS
DYNAPRIPDLFPPGQITDLNAELHSGSLILMTPTAGDGDHFAHXYILRISTSID
LRDKNESLOVNTALIPKANESEVLEFLFENGTEDGDLFAIYVAVKDLKSI
SNARVSLFIPQTPPETPSPDTSAPCPNIIHNSIIPGIHILKIMKMGELQLSIA"
exon 6667..6807
/gene="hclcal"
/number=3
6979..7126
/gene="hclcal"
/number=4
9696..9801
/gene="hclcal"
/number=5
15454..15631
/gene="hclcal"
/number=6
18591..18809
exon

```

```

/number=7
19783..20010
/gene="hclcal"
/number=8
22253..22427
/gene="hclcal"
/number=9
24522..24628
/gene="hclcal"
/number=10
26644..26859
/gene="hclcal"
/number=11
27447..27708
/gene="hclcal"
/number=12
28766..28936
/gene="hclcal"
/number=13
31840..32079
/gene="hclcal"
/number=14
32919..33522
/gene="hclcal"
/number=15
BASE COUNT 11245 a 6870 c 7230 g 9933 t
ORIGIN

```

Query Match 63.8%; Score 153; DB 9; Length 35278;
 Best Local Similarity 96.9%; Pred. No. 2,3e-37;
 Matches 156; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 4 ATGTGATGTCCTGCTGCTGAGTCTACTCTCCAGTAATGATGAACTTACAGC 63
|||||
DB 697 AGGTGATGTCCTGCTGCTGAGTCTACTCTCCAGTAATGATGAACTTACAGC 7036
|||||
QY 54 AGATGGCACTGTGAGAGAGAGGTAAGATCCACTCTGATTTCATTCAG 123
|||||
DB 7637 AGATGGCACTGTGAGAGAGAGGTAAGATCCACTCTGATTTCATTCAG 7036
|||||
QY 124 GAAAAAGTACTGTAATGACACCAAGGTAGGCAATTT 164
|||||
DB 7097 GAAAAAGTACTGTAATGACACCAAGGTAGGCAATTT 7137
|||||

```

```

RESULT 6
HSJ651E10
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
human.
Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
human.
HSJ651E10 113764 bp DNA PRI 27-MAY-2000
Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1,
complete sequence.
AL122002
AL122002.16 GI:8247274
HTG.

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
Brown, A.
Direct Submission
Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk

```

On Jun 4, 2000 this sequence version replaced gi:8247030.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence has been finished according to sequence map criteria

BASE COUNT 1098 a 594 c 633 g 879 t

ORIGIN

```

/protein_id="AA048398.1"
/db_xref="GI:5726289"
/translation="MGCFRCGVFLLVLCILHQSTSTKLNNGPEDIYVIDSVPE
DEKIIIEQLEDVTTASTYLEPATEKREFEFKNVSTLLIPNMAENQVYRPKHEKHAH
VIVAPITLPGKDEPYTKQTECEGEKEYIHETTPOLLGKQKNEGPGPKLVHEMAM
RMGVCDEFENEDQPYRAKSKRIEATRCGASGNGNRYKCGGSCLSIACIDSTKTK
YKGCQOFFPNDQVTEKASIMFQSIDSVFECNKNTHQENSLQNIKNRSTMEVY
SNSNDKILPMYTPPPPVSLIKRIQRIAYLIDMSGNGDRINRMQAKHFI
LQYENGSMQVMEHFDSPATVKNLQIKSDENITMAGIPYPLDGSTSCGIKYKA
FOVGLASHLSDGSEVYLIDGENDTASCCIDYKQSGAYHFIALGRANDENTEMK
KITGSHFYSDENQNGNLIDARCAIUSGNTDLSOKLSIQLESKULIANSAMMDYIV
IDSVGKDTFFLITWNSLPISILMDSPGITHENFTVDATSKMAYLISPGAKGTMA
YLNQAKANPETLITTVTSRANSSAPPTIVANKANNKVNNSPISYAYALAEQVYV
LGNAVTAIESQNGHTEVLELDNGSDGPKNGGVSRYPATTEMGYSKILQVRAIG
GANTARLKLRLPLNRAAYIPGMVYNGELEANRPEPEIDEDQITLQEFRSAGCAFI
VSQVPSLIDPOYPSQITDIDAVYNGELILTTAPGDNDVCKVORYIRIASITFI
DLRSPFDALQVNTTDLSPKANSKESAFENPENSIMENAHIEIATKSIDKSNLJTK
VSNIAQYTLFIQPNDDIDPTPTPTPTPKSHNSGNISLTVLSVIGSVIYVNFILTS
TTT
```

Query Match	Similarity	54.5%	Score	130.8	DB	9	Length	3204;
Best Local	Similarity	73.0%	Pred.	No.	1.6e-30;			
Matches	168:	Conservative	0;	Mismatches	62;	Indels	0;	Gaps
Oy	1	AAATGCGATGTTCTGGTGGCTGAGTCTACTCTCCACAGTAATGATGAACCTCACTG	60					
Db	324	AACATCGATGATTTATGTATGTGCACCACCTACACTCCAGTAGAGATGACCATACCA	383					
Oy	61	AGCAATGGGCACTGTGGAGAGAGGGTGAAGAAGATCCACTCACTCTGATTTTCATTC	120					
Db	384	AGCACTTACACAAAGTGGAGAGAAAGGCGAATACATTCTACCTCCCGTCACTTTCAC	443					
Oy	121	CAGAGAAAAAGTTGAGCTAATATGAGACCAACAAGTAGGAGCATTTGTCATGAGTGCTC	180					
Db	444	TTGGAAAAAACAATATGATATGAGACCAACAGCAAACTGTTTGTCCATGATGGGCTC	503					
Oy	181	ATCTACGATGGGGAGTATTTGACGAGTACAAATATGATGAGAAATTCATC	230					
Db	504	ACCTCGGTGGGGAGTGTGTGATGAGTACAAATGAGATCAGCCTTTCATC	553					
RESULT	10							
AK000072								
LOCUS								
DEFINITION								
ACCESSION								
VERSION								
KEYWORDS								
SOURCE								
ORGANISM								
REFERENCE								
AUTHORS								
TITLE								
JOURNAL								
REFERENCE								
AUTHORS								
TITLE								
JOURNAL								
COMMENT								

```

FEATURES
  source      :
               sequencing: Research Association for Biotechnology; cDNA library
               construction, 5'-6'3'-end one pass sequencing; Department of
               Virology and Human Genome Center, Institute of Medical Science,
               University of Tokyo (partly supported by Science and Technology
               Agency).
  misc_feature :
               Location/Qualifiers
               1..3221
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="COL01613"
               /clone_1lb="COL"
               /tissue_type="colon"
               /note="cloning vector pME18SFL3"
               1..3221
               /note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE
               CHANNEL PROTEIN"
BASE COUNT    1105 a      600 c      634 g      882 t
ORIGIN

```

	Query Match	54.5%	Score 130.8;	DB 9;	Length 3221;
	Best Local Similarity	73.0%;	Pred No. 1.6e-30;	Mismatches 168; Conservative	0; Indels 0; Gaps 0;
CY	1 AAAATGCAGTATTCTGGTTGCTGTACACTCCACCAGTGATGTAACCACTACACG	60			
DJ	338 AACATCGCATGTATTAGTTGCACCACTCACACTGCCAGTAGATGAACCATACCA	397			
CY	31 AGCAGATGGGCAACTGTGGAGAGAAGGGTGAAGAATCCACCTCCGATTTTCATG	120			
DJ	348 AGCAGTTCACAGAAATGTGGAGAAAAGGGGAATACATTACTTACACCCCTGACCTTCAC	457			
CY	121 CAGGAAAAAAGTTACTCTGAATATGTGACCAAGATGAGGCAATTTGTCCATGTAGTGGCTC	180			
DJ	438 TTGGAAAAAACAAATGATATGTGACCAACGGAATACTGTTGTCCATGAGTGGGCTC	517			
CY	181 ATCTACGATGGGGAGTATTGTGACGAGTACAAATGAATGAGAAATTCAC	230			
DJ	518 AACTCCGGTGGGGAGTGTTTGATGATGATCAATGAAAGATCACGCTTCTCAC	567			
RESULT .11					
LOCUS	AX092338	3265 bp	DNA		
DEFINITION	Sequence 69 from Patent WO0116318.				
ACCESSION	AX092338		PAT		21-MAR-2001
VERSION	AX092338.1	GI:13444483			
KEYWORD:	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 3265) Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A., Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Matanabe,C.K. and Wood,W.I.				
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding the same				
JOURNAL	Patent: WO 0116318-A 69 08-MAR-2001; Genentech, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..3265 /organism="Homo sapiens" /db_xref="taxon:9606"				
BASE COUNT	1159 a 596 c 632 g 878 t				
ORIGIN					
	Query Match	54.5%; Score 130.8;	DB 6;	Length 3265;	
	Best Local Similarity	73.0%;	Pred. No. 1.6e-30;	Mismatches 168; Conservative	0; Indels 0; Gaps 0;
1	AAAATGCAGTATTCTGGTTGCTGTACACTCCACCAGTGATGTAACCACTACACG	60			


```

Db 320 AACATGCGATGATATAGTTCACACCACTACACCTCCAGTAGAGATGACATACACCA 379
OY 61 AGCAGATGGGCACTGTGAGAGAGAGGCTGAAGATCCACCTGATCCATTCATTCATG 120
Db 380 AGCAGTTCACAGAAATGTGAGAGAGAGAGGCAATTCATTCATCCCTGACCTTCTAC 439
OY 121 CAGGAAAAAAGTGTGATATGATGACCAACAGTAGGCAATTTGCTCCATGAGTGGCTC 180
Db 440 TTGGAAAAAACAATGATATGAGACCAACAGTAACCTTTGCTCCATGAGTGGCTC 499
OY 181 ATTCAGATGGGAGATTTGACGAGTACATTAATGATGACAAATTTCTAC 230
Db 500 ACTCCGGTGGGAGTGTGATGATGATGACAAATGAGATGACCTTTCTAC 549

RESULT 12
AX192467 401 bp DNA PAT 15-AUG-2001
LOCUS AX192467
DEFINITION Sequence 34 from Patent WO0149716.
ACCESSION AX192467
VERSION AX192467.1 GI:15210431
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
Compounds for their use
Methods for immunotherapy and diagnosis of colon cancer and
Patent: WO 0149716-A 34 12-JUL-2001;
JOURNAL CORIXA CORPORATION (US)
FEATURES
SOURCE 1. 401
Location/Qualifiers
/db_xref="taxon:9606"
BASE COUNT 127 a 80 c 95 g 99 t
Query Match 45.8%; Score 109.8; DB 6; Length 401;
Best Local Similarity 92.9%; Pred. No. 5.1e-24;
Matches 170; Conservative 0; Mismatches 7; Indels 6; Gaps 5;

OY 1 AAAATGCTGATGTTCTGCTGGC--TGAGCTCTACTCTCCAGTAATGAGAACCTACACT 59
Db 215 AAAATGCTGATGTTCTGCTGGC--TGAGCTCTACTCTCCAGTAATGAGAACCTACACT 274
OY 60 GAGCAGAT--GGGCAACTGTGAGAGAGAGG--TGAAGGATCCACCTGACCTCTGATTTC 116
Db 275 GAGCAGATGGGGCACTGTGAGAGAGAGG--TGAAGGATCCACCTGACCTCTGATTTC 334
OY 117 ATTGACGAGAAAAAGTTCG--TGAATATGACCAACAGT--AGGCATTTGTCCATGAGT 174
Db 335 ATTGACGAGAAAAAGTTCG--TGAATATGACCAACAGT--AGGCATTTGTCCATGAGT 394
OY 175 GGG 177
Db 395 GGG 397

RESULT 13
AF043976 3415 bp mRNA PRI 07-APR-1999
LOCUS AF043976
DEFINITION Homo sapiens CLCA homolog (hCLCA3) mRNA, complete cds.
ACCESSION AF043976
VERSION AF043976.1 GI:4572288
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Gruber,A.D. and Pauli,B.U.
TITLE Molecular cloning and biochemical characterization of a truncated,
secreted member of the human family of Ca2+-activated Cl- channels
JOURNAL Biochem. Biophys. Acta 1444 (3), 418-423 (1999)
MEDLINE 99196715
REFERENCE 2 (bases 1 to 3415)
AUTHORS Gruber,A.D., Eble,R.C. and Pauli,B.U.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-1998) Department of Pathology, College of
Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES
SOURCE Location/Qualifiers
1. 3415
/db_xref="taxon:9606"
1. 3415
/gene="hCLCA3"
19. 807
/gene="hCLCA3"
/codon_start=1
/product="CLCA homolog"
/protein_id="FAD23734.1"
/db_xref="GI:4572289"
/translation="MVFSLKVLFLSLLSPVKSLSVLTLNNGYDGIYAINPSPE
DEKLIIKEMVTEASTHLPHATKORAYFERNVSLIPMTYKSEYLIPOEYDQAD
VIVADLYIKGDPTLYOCGCKGQYIHFTNPLFLNNLATYGPGRKVFVGMHL
RMGVFDEYNDQDFYISRMTTEATRCSTIETVYVMVNECKGASCIARPPRRRSQGL
YEAKCPFIPRRSOTAKSIVFMQNLDSVTEFCIEKHNKAPNL"
BASE COUNT 1149 a 625 c 630 g 1011 t
Query Match 41.5%; Score 99.6; DB 9; Length 3415;
Best Local Similarity 64.1%; Pred. No. 1.2e-20;
Matches 150; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

OY 4 ATGCTGATGTTCTGCTGGTCTGAGTCTACTCTCCAGTAATGAGAACCTACACTGAGC 63
Db 317 AGCAGATGCTCATGTTGCTGATCTCTACCTGTAATAGCAGATGATCCCTATACACTTC 376
OY 54 AGATGGCAACTGTGGAGAGAGAGAGGTGAAGATCCACCTGATCTGATTCATGAG 123
Db 377 AATATGCAATGTGGAGATTAAGGACAAATATATCATTTACTCCAAACTCTTGTGA 436
OY 124 GAAAAAAGTTCGATATGAGCAGCAGGTAAGGCAATTTGTCATGAGTGGGCTCATC 183
Db 437 CTAATTAATTGCTACCACTATGGCCTCGAGGTAAGTATTTGTCATGAGTGGGCTCATC 496
OY 134 TAGATGGGAGATTTGAGAGATGACATTAATGATGAGAAATTTCTACTTATCCA 237
Db 437 TCCGCTGGGAGATTTGATGATATATGTGACACCACTCATATATTTCCA 550

RESULT 14
BTU36445 2984 bp mRNA MAM 09-FEB-1996
LOCUS BTU36445
DEFINITION Bos taurus calcium-activated chloride channel mRNA, complete cds.
ACCESSION U36445
VERSION U36445.1 GI:1184065
KEYWORDS
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 2984)
AUTHORS Cunningham,S.A., Ayada,M.S., Bubljev,J.K., Ismailov,I.I.,
Arrate,M.P., Berdiev,B.K., Benos,D.J. and Fuller,C.M.
TITLE Cloning of an epithelial chloride channel from bovine trachea
JOURNAL J. Biol. Chem. 270 (52), 31016-31026 (1995)
MEDLINE 96125078
REFERENCE 2 (bases 1 to 2984)

```

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:06 ; Search time 16681 Seconds
(Without alignments)
175.864 Million cell updates/sec

Title: US-09-049-696-2
Perfect score: 273
Sequence: 1 GTTGCATGCAACCCCAATGT.....AGATGGCACTGTGGCAG 273

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estiba:*
7: em_estro:*
8: em_estlov:*
9: em_hic:*
10: gp_est1:*
11: gp_est2:*
12: gp_hic:*
13: gp_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	206	75.5	428	11	284017 SS284017 Po
2	182.2	66.7	745	11	BC963539 602831217
3	182.2	66.7	806	11	BC962080 602826824
4	182.2	66.7	1069	11	BC962417 602827244
5	182.2	66.7	2933	12	AK008659 Mus muscu
6	180.6	66.2	2915	12	AK007466 Mus muscu
7	170.2	62.3	767	11	BF578246 602093073
8	170.2	62.3	876	11	BF234977 602027529
9	168.8	61.8	1800	11	BF579071 602056001
10	152.4	55.8	998	11	BF583412 602097421
11	151.2	55.4	600	10	AL602665 DKFZp6860
12	127.2	46.6	485	10	BE135550 ug52g02.y

13	125.6	46.0	508	10	AI787977 u115g01.y
14	115	42.1	509	10	AM323607 uc074d01.y
15	115	42.1	568	10	BE375531 u0742541b
16	115	42.1	674	11	BE380998 601271535
17	115	42.1	761	11	BI251115 602994311
18	115	42.1	827	11	BI251115 602994311
19	115	42.1	912	11	BE913371 601665583
20	115	42.1	947	11	BE913371 601665583
21	114	41.8	1234	11	BF099622 601751643
22	113.4	41.5	519	11	BF099622 601751643
23	113.4	41.5	638	10	BF099622 601751643
24	113.4	41.5	1153	11	BF101091 601754657
25	111.8	41.0	884	11	BF102315 601751925
26	107.2	39.3	440	11	BI108239 602902275
27	104.8	38.4	759	11	BI108239 602902275
28	104.6	38.3	961	11	BI108239 602902275
29	104.6	38.3	1008	11	BI108239 602902275
30	102	37.4	813	11	BE912964 601668935
31	100.2	36.7	328	11	BE912964 601668935
32	99.6	36.5	417	10	AA245677 mx03b07.y
33	97.6	35.8	227	10	BB582380 BB582380
34	95.6	35.0	858	11	BI110624 602896268
35	94.4	34.6	585	11	BI110624 602896268
36	93.8	34.4	490	11	BI110624 602896268
37	91.2	33.4	486	11	AM767641 da84f11.y
38	90.4	33.1	607	11	BI175323 602337664
39	89.6	32.8	538	11	BI175323 602337664
40	82.6	30.3	513	11	BF230865 BF230865
41	79.4	29.1	882	10	BE288011 601093773
42	75.4	27.6	401	13	AO762814 HS_3188.A
43	75.2	27.3	629	10	BE148589 MRO-HT024
44	74.6	27.3	654	13	AZ836842 2M0131D22
45	73.6	27.0	538	10	AL595524 DKFZp311D

ALIGNMENTS

RESULT 1
284017 428 bp mRNA EST 30-JAN-1997
DEFINITION SS284017 Porcine small intestine cDNA library Sus scrofa clone
c12a03.5, mRNA sequence.

ACCESSION 284017
VERSION 284017.1 GI:1806336
KEYWORDS EST.
SOURCE Sus scrofa
ORGANISM Pig.

REFERENCE 1 (bases 1 to 428)
AUTHORS Wintero,A.K., Fredholm,M. and Davies,W.
TITLE Evaluation and characterization of a porcine small intestine cDNA library: analysis of 839 clones
JOURNAL Mamm. Genome 7 (7), 509-517 (1996)
MEDLINE 96327607

COMMENT Contact: A.K. Wintero
Department of Animal Science and Animal Health, Division of Animal
Genetics, The Royal Veterinary and Agricultural University
Bulowsvej 13, 1870 Frederiksberg C, Denmark.

FEATURES
location/Qualifiers
1..428
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="c12a03"
/clone_lib="Porcine small intestine cDNA library"
/note="directionally cloned cDNA in Xla-blue MRF"

BASE COUNT 134 a 87 c 100 g 102 t 5 others
ORIGIN

Query Match 75.5% Score 206; DB 11; Length 428;
Best local Similarity 83.8% Pred. No. 2.5e+46;

Matches 227; Conservative . 0; Mismatches 44; Indels 0; Gaps 0;

QY	2	TTGCAATGACC	CCCAATG	TGGCACA	AGTGA	AAACAT	CATTCA	CAAAAT	ATAAG	ACATGG	61
Db	135	TTGCATATTG <td>ACCCAC <td>GTGCG <td>CCAGAG <td>TGACAT <td>CTATTCA <td>GAATAT <td>ATAAG <td>ATATGG <td>194</td> </td></td></td></td></td></td></td></td>	ACCCAC <td>GTGCG <td>CCAGAG <td>TGACAT <td>CTATTCA <td>GAATAT <td>ATAAG <td>ATATGG <td>194</td> </td></td></td></td></td></td></td>	GTGCG <td>CCAGAG <td>TGACAT <td>CTATTCA <td>GAATAT <td>ATAAG <td>ATATGG <td>194</td> </td></td></td></td></td></td>	CCAGAG <td>TGACAT <td>CTATTCA <td>GAATAT <td>ATAAG <td>ATATGG <td>194</td> </td></td></td></td></td>	TGACAT <td>CTATTCA <td>GAATAT <td>ATAAG <td>ATATGG <td>194</td> </td></td></td></td>	CTATTCA <td>GAATAT <td>ATAAG <td>ATATGG <td>194</td> </td></td></td>	GAATAT <td>ATAAG <td>ATATGG <td>194</td> </td></td>	ATAAG <td>ATATGG <td>194</td> </td>	ATATGG <td>194</td>	194
QY	62	TGACCCAG <td>CGATCT <td>CTGTAT <td>CTGTTT <td>GAAGCT <td>ACAGGA <td>AAACG <td>ATTTTAT <td>TTTCAAAATG <td>121</td> </td></td></td></td></td></td></td></td>	CGATCT <td>CTGTAT <td>CTGTTT <td>GAAGCT <td>ACAGGA <td>AAACG <td>ATTTTAT <td>TTTCAAAATG <td>121</td> </td></td></td></td></td></td></td>	CTGTAT <td>CTGTTT <td>GAAGCT <td>ACAGGA <td>AAACG <td>ATTTTAT <td>TTTCAAAATG <td>121</td> </td></td></td></td></td></td>	CTGTTT <td>GAAGCT <td>ACAGGA <td>AAACG <td>ATTTTAT <td>TTTCAAAATG <td>121</td> </td></td></td></td></td>	GAAGCT <td>ACAGGA <td>AAACG <td>ATTTTAT <td>TTTCAAAATG <td>121</td> </td></td></td></td>	ACAGGA <td>AAACG <td>ATTTTAT <td>TTTCAAAATG <td>121</td> </td></td></td>	AAACG <td>ATTTTAT <td>TTTCAAAATG <td>121</td> </td></td>	ATTTTAT <td>TTTCAAAATG <td>121</td> </td>	TTTCAAAATG <td>121</td>	121
Db	195	TGACCAAA <td>GCATCT <td>CCATAT <td>CTGTTT <td>GAAAGT <td>ACAGG <td>AAAAAG <td>ATTTCT <td>ATTTCAAAATG <td>254</td> </td></td></td></td></td></td></td></td>	GCATCT <td>CCATAT <td>CTGTTT <td>GAAAGT <td>ACAGG <td>AAAAAG <td>ATTTCT <td>ATTTCAAAATG <td>254</td> </td></td></td></td></td></td></td>	CCATAT <td>CTGTTT <td>GAAAGT <td>ACAGG <td>AAAAAG <td>ATTTCT <td>ATTTCAAAATG <td>254</td> </td></td></td></td></td></td>	CTGTTT <td>GAAAGT <td>ACAGG <td>AAAAAG <td>ATTTCT <td>ATTTCAAAATG <td>254</td> </td></td></td></td></td>	GAAAGT <td>ACAGG <td>AAAAAG <td>ATTTCT <td>ATTTCAAAATG <td>254</td> </td></td></td></td>	ACAGG <td>AAAAAG <td>ATTTCT <td>ATTTCAAAATG <td>254</td> </td></td></td>	AAAAAG <td>ATTTCT <td>ATTTCAAAATG <td>254</td> </td></td>	ATTTCT <td>ATTTCAAAATG <td>254</td> </td>	ATTTCAAAATG <td>254</td>	254
QY	122	TTGCATTT <td>TGATTC <td>GTGAAC <td>ATGSA <td>AAACAT <td>AGN <td>TGAGT <td>ATGTA <td>GTAGACCAAACTTG <td>184</td> </td></td></td></td></td></td></td></td>	TGATTC <td>GTGAAC <td>ATGSA <td>AAACAT <td>AGN <td>TGAGT <td>ATGTA <td>GTAGACCAAACTTG <td>184</td> </td></td></td></td></td></td></td>	GTGAAC <td>ATGSA <td>AAACAT <td>AGN <td>TGAGT <td>ATGTA <td>GTAGACCAAACTTG <td>184</td> </td></td></td></td></td></td>	ATGSA <td>AAACAT <td>AGN <td>TGAGT <td>ATGTA <td>GTAGACCAAACTTG <td>184</td> </td></td></td></td></td>	AAACAT <td>AGN <td>TGAGT <td>ATGTA <td>GTAGACCAAACTTG <td>184</td> </td></td></td></td>	AGN <td>TGAGT <td>ATGTA <td>GTAGACCAAACTTG <td>184</td> </td></td></td>	TGAGT <td>ATGTA <td>GTAGACCAAACTTG <td>184</td> </td></td>	ATGTA <td>GTAGACCAAACTTG <td>184</td> </td>	GTAGACCAAACTTG <td>184</td>	184
Db	255	TTGCAATTT <td>TATTC <td>CTGCA <td>AGTTGG <td>NNGG <td>CCAAAA <td>CCG <td>GNATAT <td>GTGTAAACCCAAACTTG <td>314</td> </td></td></td></td></td></td></td></td>	TATTC <td>CTGCA <td>AGTTGG <td>NNGG <td>CCAAAA <td>CCG <td>GNATAT <td>GTGTAAACCCAAACTTG <td>314</td> </td></td></td></td></td></td></td>	CTGCA <td>AGTTGG <td>NNGG <td>CCAAAA <td>CCG <td>GNATAT <td>GTGTAAACCCAAACTTG <td>314</td> </td></td></td></td></td></td>	AGTTGG <td>NNGG <td>CCAAAA <td>CCG <td>GNATAT <td>GTGTAAACCCAAACTTG <td>314</td> </td></td></td></td></td>	NNGG <td>CCAAAA <td>CCG <td>GNATAT <td>GTGTAAACCCAAACTTG <td>314</td> </td></td></td></td>	CCAAAA <td>CCG <td>GNATAT <td>GTGTAAACCCAAACTTG <td>314</td> </td></td></td>	CCG <td>GNATAT <td>GTGTAAACCCAAACTTG <td>314</td> </td></td>	GNATAT <td>GTGTAAACCCAAACTTG <td>314</td> </td>	GTGTAAACCCAAACTTG <td>314</td>	314
QY	182	AGACCT <td>TACAAAA</td> <td>TGCTG</td> <td>ATTTCT</td> <td>TGGT</td> <td>TGCTGA</td> <td>GTCTANT</td> <td>CCTCC</td> <td>AGNAAATGATGAAC</td> <td>241</td>	TACAAAA	TGCTG	ATTTCT	TGGT	TGCTGA	GTCTANT	CCTCC	AGNAAATGATGAAC	241
Db	315	AGACCT <td>ACAAAA</td> <td>TGCTG</td> <td>ATTTCT</td> <td>TGGT</td> <td>ATGTA</td> <td>CAACCA</td> <td>AAATCCT</td> <td>CCAGAAAAATGATGGGC</td> <td>374</td>	ACAAAA	TGCTG	ATTTCT	TGGT	ATGTA	CAACCA	AAATCCT	CCAGAAAAATGATGGGC	374
QY	242	CCTAC <td>ACTG <td>NCAG</td> <td>ATGG <td>GCAC <td>GTGG <td>CGA</td> <td>272</td> <td></td> <td></td> </td></td></td></td>	ACTG <td>NCAG</td> <td>ATGG <td>GCAC <td>GTGG <td>CGA</td> <td>272</td> <td></td> <td></td> </td></td></td>	NCAG	ATGG <td>GCAC <td>GTGG <td>CGA</td> <td>272</td> <td></td> <td></td> </td></td>	GCAC <td>GTGG <td>CGA</td> <td>272</td> <td></td> <td></td> </td>	GTGG <td>CGA</td> <td>272</td> <td></td> <td></td>	CGA	272		
Db	375	CCTAT <td>AC <td>GAG <td>CA <td>AGAT <td>GG <td>NA <td>CT <td>NTGTG <td>AGA 405</td> </td></td></td></td></td></td></td></td>	AC <td>GAG <td>CA <td>AGAT <td>GG <td>NA <td>CT <td>NTGTG <td>AGA 405</td> </td></td></td></td></td></td></td>	GAG <td>CA <td>AGAT <td>GG <td>NA <td>CT <td>NTGTG <td>AGA 405</td> </td></td></td></td></td></td>	CA <td>AGAT <td>GG <td>NA <td>CT <td>NTGTG <td>AGA 405</td> </td></td></td></td></td>	AGAT <td>GG <td>NA <td>CT <td>NTGTG <td>AGA 405</td> </td></td></td></td>	GG <td>NA <td>CT <td>NTGTG <td>AGA 405</td> </td></td></td>	NA <td>CT <td>NTGTG <td>AGA 405</td> </td></td>	CT <td>NTGTG <td>AGA 405</td> </td>	NTGTG <td>AGA 405</td>	AGA 405

RESULT	2
LOCUS	BC963539
DEFINITION	BC963539 745 bp mRNA EST 12-JUN-2001 602831217F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4985741 5' mRNA sequence.

ACCESSION	BG963539
VERSION	BG963539.1
KEYWORDS	GI:14351176
SOURCE	EST.
ORGANISM	house mouse.
	Mus musculus

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (pages 1 to 745)	NIH-MGC	http://mgc.ncl.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
	Contact: Robert Strausberg	Ph. D.		

Email: cgabbs-r@mail.nih.gov
Tissue Procurement: jeffrey E. Green, M.D.
cdna library preparation: Life Technologies, Inc.
cdna library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnl.gov>
Plate: L1M10394 row: a column: 06
High quality sequence stop: 745.

FEATURES	Location/Qualifiers
source	1. .745

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1IMAGE:4985741"
/clone_11b="NCI_CGAP_CO24"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; cloned insertion: 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

```

Query Match	66.73;	Score 182.2;	DB 11;	Length 745;
Best Local Similarity	79.08;	Pred. No. 8.8e-40;		
Matches 214; Conservative	0;	Mismatches 57;	Indels 0;	Gaps 0

OY 2 TTGCATTCGACCCCAATGTGCCAGAGTGAACACTCATTCAACAATAAAGCATGG 61
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Dd 124 TCGCCATGACCAACAGACGTGCCGGAAGATGAAGCCCTCATTCAACAACATAAAGACATGG 189

OY	62	TGACCCAGGCGATCTCTGATCTGTTTAAAGCTACAGGAAAAGCGATTTATTTAAAAATG	121
Db	184	TGACTCAGGCGCTCTCCATACCTGTTTGAAGCTACAGGAAAAAGATTTTACTTCAAAAATG	243
OY	122	TTTGCCATTTTGATTTCTTGGAACATGAGAAGACAAAGAGTACTATGTGAGACCAAAACCTTG	181
Db	214	TTTGCCATTTTGATTTCCGAGAGCTGGAAGGCAAGCCTGATATACGAGGCCAAAACCTTG	303
OY	132	AGACCTACAAAATGCTGATGTTTGTGTTGCTGAGTCTANTCTCCAGGAAATGATGAAAC	241
Db	304	AAACCTTCAAAAAGCGATGATCTCTTTATCAACCAACACGCCCTCTTAGGCAATGATGAGC	363
OY	242	CTTACACGTGNGCAGATGAGGCAACTGTGGCGA	272
Db	364	CTTACACCGAACAATATGAGGACATGTGAGGA	394

LOCUS	BC962080	806 bp	mRNA	EST	12-JUN-2001
DEFINITION	602826824F1	NCI_CGAP_Co24	Mus musculus	CDNA clone	IMAGE:4981853
ACCESSION	BC962080				
VERSION	BC962080.1	GI:14349717			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus
1 (bases 1 to 806)
NIH-MGC <http://mgc.nhl.nih.gov/>
TITLE: National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL: Unpublished (1999)
COMMENT: Contact: Robert Strausberg, Ph.D.

Email: cgabps-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINT at:
<http://image.llnl.gov>
Plate: LLM10983 Row: o Column: 06
High quality sequence stop: 713.

```

FEATURES:
source      Location/Qualifiers
1. .806

```

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4981853"
/clone_1lb="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

```

Query Match	66.7%	Score 182.2	DB 11	Length 806
Best Local Similarity	79.0%	Pred. No. 8.8e-40		
Matches 214: Conservative	0	Mismatches 57	Indels 0	Gaps 0

Db 126 TCGGCATAGACCAAGACGTGGCGGAAGATGAAGCCCTCATTTCAACACATTAAGGACATGG 185

cy 52 TGACCCAGGCATCTCTGTATCTGTTTGAAGCTACAGGAAAGCATTTTATTCCAAAATG 121
 116 TGACTCAGGCCTCTCCATACCTGTTTGAAGCTACAGGAAAGATTTTACTTCAAAAATG 245

QY	122	TTGGCATTTTGATTTCCCTAAACATGAGAGACAAAGNACTATGATGAGACCAAACTTG	181
Db	246	TTGGCATTTTGATTTCCCGAGAGCTTGGAAAGGCAAAAGCTGAATATACGAGCCAAAACCTTG	305
QY	182	AGACCTTACAAAATGCTGATGTTCGTGGTTCGAGCTTANTCCCTCAGGMAATGATGAAC	241
Db	306	AAACCTTCAAAAAGCTGATGTTCCTTGTATATCAACAACGAGCCCTTAGGCATGATGAGC	365
QY	242	CTTACACTGNGCAGATGGCGCAACTTGGCGA	272
Db	366	CTTACACCCGAACATATAGAGAGCATGTGGAGA	396
RESULT	4		
BC962417			
LOCUS	1069 bp	mRNA	EST
DEFINITION	602827244F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4982051		12-JUN-2001
ACCESSION	BC962417		
VERSION	BC962417.1	GI:14350054	
KEYWORDS	EST.		
ORGANISM	house mouse.		
COMMENT	Mus musculus		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus. 1 (bases 1 to 1069)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
JOURNAL	Contact: Robert Strausberg, Ph.D.		
COMMENT	Email: cgapbs-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10984 Row: g Column: 12 High quality sequence stop: 586. Location/Qualifiers		
FEATURES	1..1069		
SOURCE	/organism="Mus musculus"		
	/strain="FVB/N"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:4982051"		
	/clone_lib="NCI CGAP Co24"		
	/lab_host="DH10B (TI phage-resistant)"		
	/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 kb. Constructed by life Technologies. Note: this is a NCI_CGAP library."		
BASE COUNT	297 a 397 c 219 g 156 t		
ORIGIN			
Query Match	66.7%;	Score 182.2;	DB 11; Length 1069;
Best Local Similarity	79.0%;	Pred. No. 9e-40;	
Matches	214; Conservative	0; Mismatches	57; Indels
		0; Gaps	0;
QY	2	TTGCATGACCCCAATGTCGCAGATGAAGAACACTCTTCAACAAATTAAGACATGG	61
Db	136	TGCGCATAGACACACGACGCGCGGAAGATGAAGACCCCTCTTCAACACATATAGAGACATGG	195
QY	62	TGACCCAGGCACTCTGATATCTGTTGAAGCAACGAAAGGAGATTCTTTCACAAAAG	121
Db	196	TGACTCAGGCGCTCTCATACCTGTTTGAAGCTACAGGAAAAAGATTCTTCAAAAAG	255
QY	122	TTGGCATTTTGATTTCCCTAAACATGGAAGACAAAGNACTATGAGAGACCAAACTTG	181
Db	256	TTGGCATTTTGATTTCCCGAGAGCTTGGAAAGGCAAAAGCTGAATATACGAGCCAAAACCTTG	315
QY	182	AGACCTTACAAAATGCTGATGTTCGTGGTTCGAGCTTANTCCCTCAGGMAATGATGAAC	241

Db 236 AACCTTGAAAACCGCTATGTCTTGTATCAACACACAGCCCTTAGCATGTATGAC 375

cy 242 CCTACACGTCGACATGGCAACTGTGGCA 272

Db 336 CCTACACGACACATATAGACATGTGGAGA 406

RESULT 5

AK008659 2933 bp mRNA HFC 05-JUL-2001

DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210006G11, full insert sequence.

ACCESSION AK008659

VERSION AK008659.1 GI:12842987

KEYWORDS CAP trapper.

SOURCE Mus musculus (strain:C57Bl/6J) adult male stomach cDNA to mRNA, clone:11p:RIKEN full-length enriched mouse cDNA library clone:2210006G11.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2933)

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning methods in enzymology. 303, 19-44 (1999)

JOURNAL MEDLINE 99279253

PUBMED 10349636

REFERENCE 2 (bases 1 to 2933)

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, H., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL MEDLINE 2049374

PUBMED 11042159

REFERENCE 3 (bases 1 to 2933)

AUTHORS Shibata, K., Itoh, M., Alzawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitunai, T., Tashiro, H., Itoh, M., Sumi, H., Ishii, Y., Nakamura, S., Hazama, M., Nishie, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiyake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Toned, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, K., Kita, A., and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL MEDLINE 20530913

PUBMED 11076861

REFERENCE 4 (bases 1 to 2933)

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL MEDLINE Nature 409, 685-690 (2001)

PUBMED 5 (bases 1 to 2933)

AUTHORS Adachi, J., Alzawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, Y., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL MEDLINE Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222).

/organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="4217548"
 /clone_lib="NCI-CGAP-Co24"
 /lab_host="DH10B (TI phase-resistant)"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI-CGAP library."
 BASE COUNT 316 a 238 c 250 g 194 t
 ORIGIN

Query Match 55.8%; Score 152.4; DB 11; Length 998;
 Best Local Similarity 76.8%; Pred. No. 1.3e-31;
 Matches 209; Conservative 0; Mismatches 60; Indels 3; Gaps 2;

QY 2 TTGCATCGACCCCATGTCAGAGATGAAACACTCATTCACAAATTAAGGACATG 60
 DB 120 TCGCCATGACCGACGACGCGGAGATGAAAGCCCTTCATTCAGCATAAAGGACATG 179
 QY 61 GTGACCCAGCACTCTGATCTGTTGAAGCTACAGAAAGGATTTATTTCAAAAT 120
 DB 180 GTGACTCAGGCTCTTCATCACTGTTGAAGCTACAGAAAGATTTACTTCAAAAT 239
 QY 121 GTTGCATTTGATTCCTGAAACATGAAAGAGAGTACTATGTAGACCAAACTT 180
 DB 240 GT -GCCATTTGATTCGAGAGCTGGAAGGCAAGCCGGAATATGAGAGCAAACTT 297
 QY 181 GAGACTACAAAATGCTGATGTTGCTGCTGAGTCTAATCCAGAGNAATGATGA 240
 DB 298 GAAACCTTCAAAAACGCTGATGCTTGTATCAACAACAGCAGCCCTAGGCAATGATGAG 357
 QY 241 CCTACACTGACGATGGCAACTGTGGCGA 272
 DB 358 CCTACACCGAATATGAGACATGTGAGA 389

RESULT 11
 LOCUS AL602665 600 bp mRNA EST 14-AUG-2001
 DEFINITION DKFZp68600214.t1 686 (synonym: h1cc3) Homo sapiens cDNA clone
 AL602665
 ACCESSION AL602665.1 GI:15166171
 VERSION EST.
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 600)
 AUTHORS Mambutt,R., Heubner,D., Mewes,W., Well,B. and Wiemann,S.
 TITLE EST (Mambutt,R., Heubner,D., Mewes,W., Well,B. and Wiemann,S.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Mambutt R
 MIPs
 Am Klopferspitz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No 31 sequence available.
 This clone (DKFZp68600214) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source
 1..600
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="DKFZp68600214"
 /clone_lib="686 (synonym: h1cc3)"

/tissue_type="human skeletal muscle"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiIb;
 cDNA-collection"
 BASE COUNT 211 a 115 c 120 g 154 t
 ORIGIN

Query Match 55.4%; Score 151.2; DB 10; Length 600;
 Best Local Similarity 71.7%; Pred. No. 2.7e-31;
 Matches 195; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 2 TTGCATCGACCCCATGTCAGAGATGAAACACTCATTCACAAATTAAGGACATG 61
 DB 149 TTGTTATAGATCTAGTGTCCAGAGATGAAATTAATGAAACAAATGAGAGATG 208
 QY 152 TGACCCAGCATCTCTGATCTGTTGAAGCTACAGAAAGGATTTATTTCAAAATG 121
 DB 209 TGACTCAGGCTCTTCATCACTGTTGAAGCTACAGAAAGATTTACTTCAAAAT 268
 QY 122 TTGCATTTGATTCCTGAAACATGAAAGAGTACTATGTAGACCAAACTT 181
 DB 269 TATCTATATTAATTCCTGAGAAATGGAAGAAATCTCTGATCAAAAGGCCAAACATG 328
 QY 182 AGACCTACAAAATGCTGATGTTGCTGCTGAGTCTAATCCAGAGNAATGATGA 241
 DB 329 AAAACCTAANAAGTGTGATGTTAGTTGACACACTCACTCCAGAGTATGATGAG 388
 QY 242 CCTACACTGACGATGGCAACTGTGGCGAG 273
 DB 389 CATACACCAAGCATTCACAGAAATGTGAGAG 420

RESULT 12
 LOCUS BE135550 485 bp mRNA EST 21-JUN-2000
 DEFINITION ug52g02.y1 Barstead bowel MRLB9 Mus musculus cDNA clone
 IMAGE:1546034.5' similar to TR:095151 095151 CALCIUM-DEPENDENT
 CHLORIDE CHANNEL-1. ; mRNA sequence.
 ACCESSION BE135550.1 GI:8598069
 VERSION EST.
 KEYWORDS house mouse.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 485)
 AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
 B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Rittler
 E., Kohn,S., Shu,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Maria M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@imga.llnl.gov) for further information.
 MGI:951382
 Seq primer: -40RP from Gibco
 High quality sequence stop: 466.
 FEATURES
 source
 1..485
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="1546034"
 /clone_lib="Barstead Bowel MRLB9"


```

/tissue_type="bowel"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site.1: EcoRI; Site.2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTCAGATTCGAGGAGGCGGCCCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGATCTGTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Source irradiated bowel harvested 72 hours after
irradiation (1400 Gys). Library constructed by Bob
Barstead."
BASE COUNT      154 a      102 c      108 g      121 t
ORIGIN

```

```

Query Match      46.6%; Score 127.2; DB 10; Length 485;
Best Local Similarity 66.2%; Pred. No. 1e-24;
Matches 180; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```

```

QY 2 TTGACATGACCCCAATGTGCCAGAGATGAACACATTCATCAACAATTAAGCAGATGG 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 130 TTGCATGATGCTGCTGTGCCAGAGATGAACAATTAATGAACAATTAAGGGAATGG 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 62 TGACCCAGGACATCTGTATCTGTTGAAGCTACAGGAAGGATTTATTTCAAAAATG 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 190 TGACATAAGCATCTACATACCTTTTGAAGCTACAGAAAAGATTTTCTTCAAAAATG 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 122 TTGCATTTTGAATCTCTGAACATGAGAACAGAGATGATGATGAGACCAAAATCTG 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 250 TATCTATATTAATTCCTGAGAGTGTGAAGACAGTCCCAAGAGAGGCCAAAGAG 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 132 AGACCTACAAAATGCTGATGTTGTTGCTGAGATCTAATTCCTCCAGGATGATGATGAC 241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 310 AGAGTTACAAAGCATCTGACATTAAGTGGACCTCCACGCTTGAAGGACAGATGAGC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 242 CCTACCTGNGCAGATGGGCACTGTGGCAG 273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 370 CCTACACGACGCAATTCACACAGCTGTAAGAG 401
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 13
A1787977      508 bp      mRNA      EST      02-JUL-1999
LOCUS      u115g01.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
DEFINITION IMAGE:2076912.5' similar to TR:088826 088826 GOB-5 PROTEIN. ;, mRNA
sequence.
ACCESSION A1787977
VERSION A1787977
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

```

```

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 508)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person,
B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter,
E., Korn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Other ESTs: u115g01.x1
Contact: Maria M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:993348

```

```

Seq primer: custom primer used
High quality sequence stop: 505.
Location/Qualifiers
1..508

```

```

FEATURES
Source
1..508
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone_image="IMAGE:2076912"
/clone_lib="Sugano mouse embryo mewa"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B"
/note="Vector: pME18S-FL3; Site.1: DraIII (CACTGCTG);
Site.2: DraIII (CACCAGTGT); 1st strand cDNA was primed
with an oligo(dT) primer [ATGCGCCCTTTTCTTTTCTTTTCTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[CTTGCCTCTTAAAGCTGGG and 3' end primer
CTTGCCTCTTAAAGCTGGG and 3' end primer
CGACCTGCACGCTGACGACA."
BASE COUNT      160 a      104 c      116 g      127 t      1 others
ORIGIN

```

```

Query Match      46.0%; Score 125.6; DB 10; Length 508;
Best Local Similarity 65.8%; Pred. No. 2.e-24;
Matches 179; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

```

```

QY 2 TTGACATGACCCCAATGTGCCAGAGATGAACACATTCATCAACAATTAAGCAGATGG 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 142 TTGCATGATGCTGCTGTGCCAGAGATGAACAATTAATGAACAATTAAGGGAATGG 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 62 TGACCCAGGACATCTGTATCTGTTGAAGCTACAGGAAGGATTTATTTCAAAAATG 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 202 TGACATAAGCATCTACATACCTTTTGAAGCTACAGAAAAGATTTTCTTCAAAAATG 261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 132 TTGCATTTTGAATCTCTGAACATGAGAACAGATGATGATGAGACCAAAATCTG 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 252 TGTCTATATTAATTCCTGAGAGTGTGAAGAACAGTCCCAATGACAGAAAGCAAG 321
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 132 AGACCTACAAAATGCTGATGTTGTTGCTGAGATCTAATTCCTCCAGGATGATGATGAC 241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 332 AGAGTTACAAAGCATCTGACATTAAGTGGACCTCCCGCTTGAGGAGCAGAGATGAGC 381
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 212 CCTACCTGNGCAGATGGGCAACTGTGGCAG 273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 382 CCTACACGACGCAATTCACACAGCTGTGAAGAG 413
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 14
AM323607      509 bp      mRNA      EST      26-JAN-2000
LOCUS      uc074d01.y1 NCI CGAP Maml Mus musculus cDNA clone IMAGE:2646257.5'
DEFINITION similar to TR:088860 088860 CHLORIDE CHANNEL CACC. ;, mRNA
sequence.
ACCESSION AM323607
VERSION AM323607
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 509)
NCI-CGAP http://www.nci.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: uc074d01.x1
Contact: Robert Strausberg, Ph.D.

```


This Page Blank (uspto)

XX Nakanishi A, Morita S;
XX WP1: 2001-355935/37.
XX P-PSDB: AAB73716.
XX
XX New antisense nucleotide, useful for treatment and prevention of
XX bronchial asthma and chronic obstructive pulmonary disease -
XX
XX Claim 3; Page 82-84; 104pp: Japanese.
XX
XX The invention relates to an antisense nucleotide targeted to the mouse
XX Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
XX the C1CA1 gene (coding sequence shown in AAH46102). The invention also
XX relates to an antibody specific for the Gob-5 protein, medical and
XX diagnostic compositions containing the antisense nucleotide or the
XX antibody, and methods and kits for screening for compounds which inhibit
XX the protein. Gob-5 and C1CA1 are proteins expressed by goblet cells.
XX The antisense oligonucleotides and antibody are therefore useful for the
XX treatment and prevention of bronchial asthma and chronic obstructive
XX pulmonary disease. The present sequence represents the human C1CA1
XX gene coding sequence.
XX
SQ Sequence 2742 BP; 833 A; 616 C; 623 G; 670 T; 0 other;

Query Match 100.0%; Score 218; DB 22; Length 2742;
Best Local Similarity 100.0%; Pred. No. 9.8e-62;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 GAACAATGGCTCATGATGATCTTTGGGCGCTTTCATCAGAAATGAGCTGTCTCA 60
DB 1392 gaacaatggctcatgattgatttggggcgttctcattcagaagaatgagctgtctca 1451
XX
XX 61 GGGCTCCATCCAGCTTGAGAGTAAGGATTAACCCCTCCAGAACGCGATGATGAATG 120
DB 1452 ggcctccatccagcttgagagtaaggattaaacctccagaaagcagctgagaa 1511
XX
XX 121 CACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
DB 1512 caagtgatgattgattgattgattgattgattgattgattgattgattgattgatt 1571
XX
XX 181 GCAGCTCCCAAAATCTTCTGTGGATCCCACTGATGATGATGATGATGATGATG 218
DB 1572 gcagctcccaaatcttctgtggaatcccaactgtggtggtggtggtggtggtg 1609
XX

RESULT 2
AAF81927
ID AAF81927 standard; cDNA: 2745 BP.
XX
XX AAF81927;
XX
XX 13-JUN-2001 (first entry)
XX
XX Human ICAC-1 nucleotide sequence.
XX
XX ICAC-1; ICAC-2; asthma; atopic allergy; asthma-related disorder;
XX interleukin 9 induced calcium activated chloride channel; IL-9;
XX calcium activated chloride channel; anti-allergic; anti-asthmatic;
XX anti-inflammatory; immunomodulatory; cystic fibrosis;
XX inflammatory bowel disease; autoimmune disease; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..2745
XX /ftag- a
XX /product= "ICAC-1"
XX /note= "IL-9 induced calcium activated chloride channel"
XX
XX MO9944620-A1.

PD 10-SEP-1999.
XX
XX 03-MAR-1999: 99WO-US04703.
XX
XX 03-MAR-1998: 98US-0076815.
XX
XX (MAGA-) MAGANIN PHARM INC.
XX
XX Holroyd KJ, Levitt RC, Maloy WL, Louhed J, McLane M;
XX Nicolides NC, Zhou Y, Dong O;
XX WP1: 1999-550979/46.
XX P-PSDB: AAB74824.
XX
XX New nucleic acid encoding calcium activated chloride channel, used to
XX identify, e.g. specific modulators for treating atopic allergy -
XX
XX Claim 1; Fig 4B; 75pp: English.
XX
XX The present sequence encodes the human interleukin 9 (IL-9) induced
XX calcium activated chloride channel 1 (ICAC-1) protein. ICAC proteins
XX have anti-allergic, anti-asthmatic, anti-inflammatory and
XX immunomodulatory activities. Compounds (A) that downregulate ICAC are
XX used to alleviate asthma (or more generally atopic allergy), while those
XX (B) that upregulate ICAC are used to treat cystic fibrosis (CF) and
XX inflammatory bowel disease (IBD) (or other autoimmune diseases).
XX Measuring levels of ICAC-1 can be used in the diagnosis of asthma
XX (increased levels) or IBD (reduced levels), also for monitoring
XX treatment of these conditions. The ICAC proteins can be used:
XX (1) to raise specific antibodies (Ab), useful: (a) as immunoassay
XX reagents, and (b) as therapeutic (A); (11) as fragments or mutants, as
XX (A) to disrupt bonding between ICAC and its ligands, and (111) to
XX identify modulators and binding partners. ICAC polynucleotides can be
XX used to generate transgenic animals or recombinant cells, used to screen
XX for antagonists, also as a source of therapeutic antisense agents or
XX diagnostic probes (for quantifying mRNA expression, e.g. for
XX identification of modulators).
XX
SQ Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;

Query Match 100.0%; Score 218; DB 20; Length 2745;
Best Local Similarity 100.0%; Pred. No. 9.8e-62;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 GAACAATGGCTCATGATGATCTTTGGGCGCTTTCATCAGAAATGAGCTGTCTCA 60
DB 1392 gaacaatggctcatgattgatttggggcgttctcattcagaagaatgagctgtctca 1451
XX
XX 61 GGGCTCCATCCAGCTTGAGAGTAAGGATTAACCCCTCCAGAACGCGATGATGAATG 120
DB 1452 ggcctccatccagcttgagagtaaggattaaacctccagaaagcagctgagaa 1511
XX
XX 121 CACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
DB 1512 caagtgatgattgattgattgattgattgattgattgattgattgattgattgatt 1571
XX
XX 191 GCAGCTCCCAAAATCTTCTGTGGATCCCACTGATGATGATGATGATGATGATG 218
DB 1572 gcagctcccaaatcttctgtggaatcccaactgtggtggtggtggtggtggtg 1609
XX

RESULT 3
AAH46124
ID AAH46124 standard; cDNA: 2825 BP.
XX
XX AAH46124;
XX
XX 11-SEP-2001 (first entry)
XX
XX Human C1CA1 cDNA, SEQ ID NO:26.
XX
XX Human C1CA1; goblet cell; mouse Gob-5 orthologue; drug screening;
XX

KW	expression inhibition; antisense therapy; gene therapy;
KM	chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
KX	ss.
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	23..2767
CDS	
FT	/tag= a
FT	/product= "Human CLCA1"
FT	/transl_except= (pos:476..478, aa:Lys)
PN	MO200138530-A1.
XX	
PD	31-MAY-2001.
XX	
PE	22-NOV-2000; 2000WO-JP08232.
XX	
PR	24-NOV-1999; 99JP-0333479.
PR	27-APR-2000; 2000JP-0127589.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
PI	Nakanishi A, Morita S:
XX	
DR	WPI: 2001-355935/37.
DR	P-PSDB: AAB73716.
XX	
PT	New antisense nucleotide, useful for treatment and prevention of
PT	bronchial asthma and chronic obstructive pulmonary disease -
XX	
PS	Example 5; Page 92-94; 104pp; Japanese.
CC	The invention relates to an antisense nucleotide targeted to the mouse
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC	the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC	relates to an antibody specific for the Gob-5 protein, medical and
CC	diagnostic compositions containing the antisense nucleotide or the
CC	antibody, and methods and kits for screening for compounds which inhibit
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC	The antisense oligonucleotides and antibody are therefore useful for the
CC	treatment and prevention of bronchial asthma and chronic obstructive
CC	pulmonary disease. The present sequence represents human CLCA1 cDNA.
XX	
SQ	Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;
Query Match	100.0%; Score 218; DB 22; Length 2825;
Best Local Similarity	100.0%; Pred. No. 1e-61;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0.	
OY	1 GAACAATGAGCCTCATTTGAATGTCTTGGGGCCCTTCATCAGAAATGAGCTGTCTCA 60
Db	1414 gaacaatgagcctcatgtatgtcttggggcccttcatacagaatgaatgagcgctgtctca 1473
OY	61 GCGETCATCCACGTTGAGAGTAGGAGATTAAACCTCCAGAACAGCACCGATGATGATG 120
Db	1474 ggcctccatccacgcttgagagtaaggattaacccctcagaaacagccagtgtgatgag 1533
OY	121 CACAGTGATCGGTGGAGCAGCACCGGTGGGAAAAGAACACTTTGTTTATACCTGGACAC 180
Db	1534 cacagtgtatcgtlgacagcacgcgtggaaaagacaccttgtttctatcactcgtgacaac 1593
OY	181 GCAGCGTCCCCAATAATCTTCTCTGGAGGCCAGTGGAC 218
Db	1594 gcagcgtcccccaaatcctctctctggaagaccagtgagac 1631
RESULT	4
ID	AAH34879
CD	AAH34879 standard; cDNA: 2854 BP.
CC	AAH34879;

```

XX 13 Human colon cancer antigen encoding CDNA SEQ ID NO:1961.
DE XX 03-SEP-2001 (first entry)
XX
XX Human colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma; chromosome 1; ss.
XX Homo sapiens,
GS WO200122920-A2.
XX
XX 28-SEP-2000; 2000WO-US26524.
XX PF
XX 29-SEP-1999; 99US-0157137.
XX PR
XX 03-NOV-1999; 99US-0163280.
XX PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruven SM, Barash SC, Birse CE, Rosen CA;
XX
XX WPI: 2001-235357/24.
XX ER
XX P-SDB; AAG73474.
XX DR
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1: Page 3462-3463; 9803pp; English.
PS
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAH77789 represent sequences used in the exemplification of the
CC present invention.
CC R.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other:
XX
XX Query Match 100.0%; Score 218; DB 22; Length 2854;
XX Best Local Similarity 100.0%; Pred. No. 1e-61;
XX Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```


RESULT 5
AAAF81787 standard; cDNA: 2854 BP.
ID AAF81787
XX AAF81787;
AC AAF81787;
XX AAF81787;
DT 12-JUN-2001 (first entry)
XX
XX Human secreted protein gene 1 SEQ ID NO:11.
DE
XX Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW immunostimulant; cytoskeletal; cardiant; vascular; anti-angiogenic;
KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
KW antialzheimer; antiparkinsonian; antimicrobial; vulnary; gene therapy;
KW immune disorder; hyperproliferative disorder; cardiovascular disease;
KW cancer; angiogenic disorder; neurological disorder; infectious disease;
KW wound healing; regeneration; chemotaxis; chromosome 1; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200112775-A2.
PN
XX
XX 22-FEB-2001.
PD
XX
XX 16-AUG-2000; 2000WO-US22325.
PE
XX
XX 17-AUG-1999; 99US-0149182.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;
PI Blise CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;
PI
XX WPI: 2001-147550/15.
DR
XX P-PSDB; AAB74733.
DR
XX
XX Nucleic acids encoding 25 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
PT
XX
XX Claim 1; Page 441; 485pp; English.
PS
XX
XX AAF81787 to AAF81817 encode the human secreted proteins given in AAB74733
CC to AAB74772. Human secreted proteins can have activities based on the
CC tissues and cells they are expressed in. Example of activities include:
CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;
CC antiinflammatory; anti-HIV; immunostimulant; cytoskeletal; cardiant;
CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
CC anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and
CC vulnary. Human secreted proteins can be used in gene therapy and
CC vaccine. Human secreted protein nucleotide sequences (NMI) and proteins
CC (PEPI) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. For example, NMI
CC and PEPI may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patients genome
CC that affect the activity of proteins by expressing inactive proteins or
CC to supplement the patients own production of polypeptides. Disorders that
CC may be prevented, diagnosed and/or treated include immune disorders,
CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,
CC angiogenic disorders, neurological disorders, infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to
CC AAF81786 and AAB74732 represent sequences used in the exemplification of
CC the present invention.
XX
XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;

Query Match 100.0%; Score 218; DB 22; Length 2854;
Best Local Similarity 100.0%; Pred. No. 1e-61;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y I GAACAAATGGCCATATGATGCTTTTGGGGCCCTTTCATCAGGAATGGAGCTGCTCTCA 60
DB 1426 gaacaatggccatattgaatgcttttggggcccttcatcaggaatggagctgctctca 1485
0Y 61 GCGCTCCATCCAGCTTGAGAGTAAGGATTAACCTTCACAGAACGCCAGTGAATGG 120
DB 1486 gcgctccatccagcttgagagtaaggattaaccttcacagaaagccagtgatgaatgg 1545
0Y 121 CACAGTATGCTGGACAGCACCGTGGGAAAGACACTTTTCTTATCCTCGACAC 180
DB 1546 cacagtatgcttgagagcacccgtggaaagacacttttcttataccctcgagaaac 1605
0Y 181 GCAGCTCCGCCAAATCCTTCCTGGGATGCCAGTGAC 218
DB 1606 gcagctccgccaaatccttctcgggagtcagtgac 1643
RESULT 6
AAH33285 standard; cDNA: 2867 BP.
ID AAH33285
XX
XX AAH33285;
AC
XX
XX 03-SEP-2001 (first entry)
DT
XX
XX Human colon cancer antigen encoding cDNA SEQ ID NO:341.
DE
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 1; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200122920-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 28-SEP-2000; 2000WO-US26524.
PE
XX
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Ruben SM, Barash SC, Blise CE, Rosen CA;
PI
XX
XX WPI: 2001-235357/24.
DR
XX P-PSDB; AAG73854.
DR
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
PT
XX
XX Claim 1; Page 2452-2453; 9803pp; English.
PS
XX
XX AA332943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytoskeletal activity and can be used in gene
CC library and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 2867 BP; 891 A; 633 C; 643 G; 700 T; 0 other;
SQ

Query Match 100.0%; Score 218; DB 22; Length 2867;
Best Local Similarity 100.0%; Pred. No. 1e-61;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAATGGCCATCATGATGCTTTTGGGGCCCTTTCATCAGAAATGAGACTGCTCTCA 60
DB 1430 gaacaatggccatcatgattgcttttggggcccttctacagaaatgagactgctctca 1489
QY 61 GCGCTCCATCAGCTTGAGAGTAAGGATTAACTCCAGACAGCCAGTGATGAATGG 120
DB 1490 gcgctccatcagcttgaagagtaaggattaaacctccagaaagcagtgatgaatgg 1549
QY 121 CACAGTATGATGTGACAGCAGCGTGGGAAGAGACATTTTCTTATCAGCTGGAGAAC 180
DB 1550 cacagtatgatgtgacagcagcgagggaagagacatttcttatacaccugaacac 1609
QY 181 GCAGCTCCCAATCCTTCTGTGGATCCAGTGCAC 218
DB 1610 gcagctcccaatccttctgtggatccagtgac 1647

RESULT 7
AAH35019
ID AAH35019 standard; cDNA; 3109 BP.

XX AAH35019;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:2101.

XX Human: colon cancer; colon cancer antigen; diagnosis; detection;

KW colorectal carcinoma; ss.

XX Homo sapiens.

PN WO200122920-A2.

XX 05-APR-2001.

PE 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI: 2001-235357/24.

DR P-PSDB; AAG35614.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX

PS Claim 1; Page 3587-3588; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing or
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell

CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAH7789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

SQ Sequence 3109 BP; 970 A; 652 C; 687 G; 797 T; 3 other;

Query Match 100.0%; Score 218; DB 22; Length 3109;
Best Local Similarity 100.0%; Pred. No. 1e-61;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAATGGCCATCATGATGCTTTTGGGGCCCTTTCATCAGAAATGAGACTGCTCTCA 60
DB 1279 gaacaatggccatcatgattgcttttggggcccttctacagaaatgagactgctctca 1338
QY 61 GCGCTCCATCAGCTTGAGAGTAAGGATTAACTCCAGACAGCCAGTGATGAATGG 120
DB 1339 gcgctccatcagcttgaagagtaaggattaaacctccagaaagcagtgatgaatgg 1398
QY 121 CACAGTATGATGTGACAGCAGCGTGGGAAGAGACATTTTCTTATCAGCTGGAGAAC 180
DB 1499 cacagtatgatgtgacagcagcgagggaagagacatttcttatacaccugaacac 1458
QY 181 GCAGCTCCCAATCCTTCTGTGGATCCAGTGCAC 218
DB 1459 gcagctcccaatccttctgtggatccagtgac 1496

RESULT 8

AAZ09840
ID AAZ09840 standard; cDNA; 3111 BP.

XX AAZ09840;

DE 26-NOV-1999 (first entry)

DE Human membrane spanning protein MSP-5 cDNA fragment 2.

KW Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;
KW neoplastic disorder; immunological disorder; reproductive disorder;

KW MSP-5; ds.

XX Homo sapiens;

PN WO946380-A2;

PD 16-SEP-1999.

PE 05-MAR-1999; 99WO-US05073.

PR 13-MAR-1998; 98US-0039064.

PA (INCY-) INCYTE PHARM INC.

PI Tang YT, Bandman O, Lal P, Hillman JL, Yue H, Corley NC;

DR P-PSDB; AAY33298.

XX New human membrane spanning proteins used to, e.g. prevent and treat
XX neoplastic disorders -
XX

PS Example 1; Page 80-81; 81pp; English.

XX This invention describes novel human membrane spanning proteins (MSPs),
CC and the polynucleotides encoding them. The products of the invention are
CC used to diagnose, prevent and treat neoplastic, immunological and
CC reproductive disorders. This sequence encodes a human membrane spanning

CC protein MSP-5 fragment.
 XX Sequence 3111 BP; 968 A; 664 C; 691 G; 788 T; 0 other;

Query Match 100.0%; Score 218; DB 20; Length 3111;
 Best Local Similarity 100.0%; Pred. No. 1e-61;
 Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAACAATGAGCCCTCATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGAGCTGCTCTCA 60
 1425 gaacaatgagccctcatgattgctttgggcgcccttccatcagaagaatgagagctgctctca 1484
 61 GCGCTCATCCAGCTTGAGAGTAAGGATTAACCTCCAGAACACCCAGTGGATGAATG 120
 1485 ggcctcatccagcttgagagtaaggaatgaacccctccagacagccagtgatgaatg 1544
 121 CACAGTGAATCGTGAGACAGCAGCGTGGAAGAGACACTTTGTTTATCAGCTGGACAA 180
 1545 cacagtgaatcgtgagacagcagcgctggaaagacacttgtttctatcactgagacac 1604
 181 GCAGCCTCCCAATCTCTCTGGATGCCAGTGGAC 218
 1605 gcagcctcccaatctctctctggatgccagtgagac 1642

RESULT 9
 AAI29502
 ID AAI29502 standard; cDNA; 3311 BP.

AAI29502;
 12-OCT-2001 (first entry)

C902P determined cDNA sequence.

Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 Immunogenic; gene therapy; vaccine; colonic cancer; ss.

Homo sapiens.

MO200149716-A2.

12-JUL-2001.

29-DEC-2000; 2000WO-US35596.

30-DEC-1999; 99US-0476296.

10-JAN-2000; 2000US-0480321.

15-FEB-2000; 2000US-0504629.

06-MAR-2000; 2000US-0519444.

19-MAY-2000; 2000US-0575251.

29-JUN-2000; 2000US-0609448.

28-AUG-2000; 2000US-0649811.

(CORI-) CORIXA CORP.

Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;

King GE, Wang T, Jiang Y;

WPI: 2001-441847/47.

Colon tumor associated proteins and nucleic acids useful for the

prevention, diagnosis and treatment of colonic cancer -

Claim 2; Page 425-426; 472pp; English.

The present invention describes colon tumour associated proteins (I) and

the polynucleotides (II) that encode them. (I) have cytostatic activity.

(I) and (II) can be used in gene therapy and vaccine production. (I) and

(II) may be used in the prevention, diagnosis and treatment of diseases

associated with inappropriate colon tumour associated protein (TCAP)

expression, such as colonic cancer. For example, (I) and (II) may be

used to treat disorders associated with decreased expression by
 rectifying mutations or deletions in a patient's genome that affect the
 activity of TCAPs by expressing inactive proteins or to supplement the
 patients own production of them. Additionally, (II) may be used to
 produce the TCAP proteins, by inserting the nucleic acids into a host
 cell culturing the cell to express the protein. (II) and its
 complementary sequences may also be used as DNA probes in diagnostic
 polymerase chain reaction (PCR) and hybridisation assays to detect and
 quantitate the presence of similar nucleic acids in samples, and
 therefore which patients may be in need of restorative therapy. (I) may
 also be used as antigens in the production of antibodies against TCAPs
 and in assays to identify modulators of TCAP expression and activity.
 Anti-(I) antibodies and antagonists may also be used to down regulate
 TCAP expression and activity. The anti-(I) antibodies may also be used
 as diagnostic agents for detecting the presence of TCAPs in samples
 (e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
 and AAI24494 to AAI24523 represent nucleotide and amino acid sequences
 given in the exemplification of the present invention.

Query Match 100.0%; Score 218; DB 22; Length 3311;
 Best Local Similarity 100.0%; Pred. No. 1e-61;
 Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAACAATGAGCCCTCATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGAGCTGCTCTCA 60
 1743 gaacaatgagccctcatgattgctttgggcgcccttccatcagaagaatgagagctgctctca 1802
 61 GCGCTCATCCAGCTTGAGAGTAAGGATTAACCTCCAGAACACCCAGTGGATGAATG 120
 1803 ggcctcatccagcttgagagtaaggaatgaacccctccagacagccagtgatgaatg 1862
 121 CACAGTGAATCGTGAGACAGCAGCGTGGAAGAGACACTTTGTTTATCAGCTGGACAA 180
 1863 cacagtgaatcgtgagacagcagcgctggaaagacacttgtttctatcactgagacac 1922
 181 GCAGCCTCCCAATCTCTCTGGATGCCAGTGGAC 218
 1923 gcagcctcccaatctctctctggatgccagtgagac 1960

RESULT 10
 AAI46101
 ID AAI46101 standard; DNA; 2739 BP.

AAI46101;

11-SEP-2001 (first entry)

Mouse Gob-5 coding sequence, SEQ ID NO:3.

Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;

expression inhibition; antisense therapy; gene therapy; bronchial asthma;

chronic obstructive pulmonary disease; antiasthmatic; ds.

Mus sp.

Key: location/qualifiers

ET CDS: 1..2739

ET CDS: 1..2739

ET CDS: 1..2739

ET CDS: 1..2739

ET CDS: 1..2739

ET CDS: 1..2739

ET CDS: 1..2739

ET CDS: 1..2739

27-APR-2000; 2000JP-0127589.
 (TAKE) TAKEDA CHEM IND LTD.
 Nakanishi A, Morita S;
 WPI, 2001-355935/37.
 P-PSDB; AAB73715.
 New antisense nucleotide, useful for treatment and prevention of
 bronchial asthma and chronic obstructive pulmonary disease -
 Claim 3; Page 80-82; 104pp; Japanese.
 The invention relates to an antisense nucleotide targetted to the mouse
 Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 the C1CA1 gene (coding sequence shown in AAH46102). The invention also
 relates to an antibody specific for the Gob-5 protein, medical and
 diagnostic compositions containing the antisense nucleotide or the
 antibody, and methods and kits for screening for compounds which inhibit
 the protein. Gob-5 and C1CA1 are proteins expressed by goblet cells.
 The antisense oligonucleotides and antibody are therefore useful for the
 treatment and prevention of bronchial asthma and chronic obstructive
 pulmonary disease. The present sequence represents the mouse Gob-5
 gene coding sequence.
 Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other;
 SQ
 Query Match 71.7%; Score 156.2; DB 22; Length 2739;
 Best Local Similarity 82.5%; Pred. No. 2e-41;
 Matches 179; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
 QY 1 GAACAATGGCCATCATGATGCTTTGGGGCCCTTCATCAGAAATGAGCTGTCTCA 60
 Db 1395 gaacaatgctctgttgcagcttcgcagcactccctccggaagcgcgcgcgcgcgc 1454
 QY 61 GCGCTCCATCCAGCTTGAGAGTAAGGATTTAACCTCCAGAACAGCCAGTGAATGG 120
 Db 1455 gactccatccagctgagagcagggagtaatcccaagaatacgaatgagatgg 1514
 QY 121 CACAGTATCGTGGAGACAGCCCTGGGAAAGACACTTGTCTTATCACCCTGACAC 180
 Db 1515 ctcaatgacgtgagacagctcggtggcaagcacttcttctatcacctggacaac 1574
 QY 181 GCAGCTCCCAAAATCCTTCTCTGGGATCCAGTGA 217
 Db 1575 gcatcctctacataattatcttgagatcccgagcga 1611
 RESULT 11
 AAH46120
 ID AAH46120 standard; cDNA; 2843 BP.
 AC AAH46120;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Mouse Gob-5 cDNA, SEQ ID NO:22.
 XX
 KW Mouse Gob-5; murine; goblet cell; human C1CA1 orthologue; drug screening;
 KW expression inhibition; antisense therapy; gene therapy; bronchial asthma;
 KW chronic obstructive pulmonary disease; antiasthmatic; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 15..2756
 FT /tag= a
 FT /product= "Mouse Gob-5"
 XX
 PN WO200138530-A1.

30 MAY-2001.
 22 NOV-2000; 2000MO-JP08232.
 24 NOV-1999; 99JP-0333479.
 27 APR-2000; 2000JP-0127589.
 (TAKE) TAKEDA CHEM IND LTD.
 Nakanishi A, Morita S;
 WPI, 2001-355935/37.
 P-PSDB; AAB73715.
 New antisense nucleotide, useful for treatment and prevention of
 bronchial asthma and chronic obstructive pulmonary disease -
 Example 1; Page 89-91; 104pp; Japanese.
 The invention relates to an antisense nucleotide targetted to the mouse
 Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 the C1CA1 gene (coding sequence shown in AAH46102). The invention also
 relates to an antibody specific for the Gob-5 protein, medical and
 diagnostic compositions containing the antisense nucleotide or the
 antibody, and methods and kits for screening for compounds which inhibit
 the protein. Gob-5 and C1CA1 are proteins expressed by goblet cells.
 The antisense oligonucleotides and antibody are therefore useful for the
 treatment and prevention of bronchial asthma and chronic obstructive
 pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
 Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;
 SQ
 Query Match 71.7%; Score 156.2; DB 22; Length 2843;
 Best Local Similarity 82.5%; Pred. No. 2e-41;
 Matches 179; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
 QY 1 GAACAATGGCCATCATGATGCTTTGGGGCCCTTCATCAGAAATGAGCTGTCTCA 60
 Db 1469 gaacaatgctctgttgcagcttcgcagcactccctccggaagcgcgcgcgcgcgc 1468
 QY 61 GCGCTCCATCCAGCTTGAGAGTAAGGATTTAACCTCCAGAACAGCCAGTGAATGG 120
 Db 1469 gactccatccagctgagagcagggagtaatcccaagaatacgaatgagatgg 1528
 QY 121 CACAGTATCGTGGAGACAGCCCTGGGAAAGACACTTGTCTTATCACCCTGACAC 180
 Db 1529 ctcaatgacgtgagacagctcggtggcaagcacttcttctatcacctggacaac 1588
 QY 181 GCAGCTCCCAAAATCCTTCTCTGGGATCCAGTGA 217
 Db 1529 gcatcctctacataattatcttgagatcccgagcga 1625
 RESULT 12
 AAF81925
 ID AAF81925 standard; cDNA; 2931 BP.
 AC AAF81925;
 XX
 DT 13-JUN-2001 (first entry)
 XX
 DE Murine ICACC-1 nucleotide sequence.
 XX
 KW ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
 KW interleukin 9 induced calcium activated chloride channel; IL-9;
 KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
 KW anti-inflammatory; immunomodulatory; cystic fibrosis;
 KW inflammatory bowel disease; autoimmune disease; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 15..2756
 FT /tag= a
 FT /product= "Mouse Gob-5"
 XX
 PN WO200138530-A1.

FT CDS 8.2749
 FT /tag= a
 FT /product= "ICACC-1"
 FT /note= "IL-9 induced calcium activated chloride channel"
 PN MO9944620-A1.
 XX 10-SEP-1999.
 PD 03-MAR-1999; 99WO-US04703.
 XX 03-MAR-1999; 98US-0076815.
 PR 03-MAR-1998; 98US-0076815.
 XX
 XX (MAG-1) MAGATININ PHARM INC.
 XX
 PI Holroyd KJ, Levitt RC, Maloy WL, Louahed J, McLane M;
 PI Nicolaides NC, Zhou Y, Dong Q;
 XX WPI: 1999-550979/46.
 DR P-PSDB; AAB74822.
 XX
 XX New nucleic acid encoding calcium activated chloride channel, used to
 PT identify, e.g. specific modulators for treating atopic allergy -
 XX
 PS Claim 2; Fig 2; 75pp; English.
 CC The present sequence encodes the murine interleukin 9 (IL-9) induced
 CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
 CC have anti-allergic, anti-asthmatic, anti-inflammatory and
 CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
 CC used to alleviate asthma (or more generally atopic allergy), while those
 CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
 CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
 CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma
 CC (increased levels) or IBD (reduced levels), also for monitoring
 CC treatment of these conditions. The ICACC proteins can be used:
 CC (1) to raise specific antibodies (Ab), useful: (a) as immunoassay
 CC reagents, and (b) as therapeutic (A), (11) as fragments or mutants, as
 CC (A) to disrupt bonding between ICACC and its ligands; and (11) to
 CC identify modulators and binding partners. ICACC polynucleotides can be
 CC used to generate transgenic animals or recombinant cells, used to screen
 CC for antagonists, also as a source of therapeutic antisense agents or
 CC diagnostic probes (for quantifying mRNA expression, e.g. for
 CC identification of modulators).
 XX
 SQ Sequence 2931 BP; 867 A; 719 C; 684 G; 661 T; 0 other;
 Query Match 71.7%; Score 156.2; DB 20; Length 2931;
 Best Local Similarity 82.5%; Pred. No. 2e-41; Mismatches 38; Indels 0; Gaps 0;
 Matches 179; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
 QY 1 GAACAATGGCCATCATGATGCTTTGGGGCCCTTTCATCAGGAAGATGAGCTGTCTCA 60
 DB 1402 gaacaatggccatcatgctttggggccctttcacaggaagatgagctgtctca 1461
 QY 61 GCGCTCCATCCAGCTTGAGAGTAAGGATTAACCTCCAGAACACCCAGTGAATGG 120
 DB 1462 gcgctccatccagcttgagagtaaggattaacctccagaacacccagtgaatgg 1521
 QY 121 CACAGTGAATCGGAGACAGCAGCGTGGGAAGAGACACTTGTCTTATCAGCTGGCAAC 180
 DB 1522 ctacagtgaatcggagacagcagctgggaagagacacttgtcttattcagctggcaac 1581
 QY 181 GCAGCCTCCCAAAATCCTTCTCTGGAGATCCAGTGA 217
 DB 1582 gcagcctcccaaaatccttctctggagatccagtga 1618
 RESULT 13
 AAA64335
 ID AAA64335 standard; DNA; 2616 BP.
 XX

AC AAA64335;
 XX 20-DEC-2000 (first entry)
 DE Clone 251688 of a novel gene associated with colon disease.
 XX
 XX Colon disease; colon cancer gene; carbonic anhydrase; galectin;
 KW carcinoembryonic antigen; colorectal carcinoma tumour associated antigen;
 KW fatty-acid binding protein; glutathione peroxidase; guanylin;
 KW cyclokeratin; cadherin; intestinal mucin; colon cancer; Crohn's disease;
 KW metastatic colon cancer; atrophic gastritis; cholecystitis;
 KW irritable bowel syndrome; ulcerative colitis; ss.
 XX
 OS Homo sapiens;
 XX
 XX WO2000050588-A2.
 XX
 XX 31-AUG-2000.
 XX
 XX 01-FEB-2000; 2000WO-US02595.
 XX
 XX 22-FEB-1999; 99US-0255381.
 XX
 XX (JNCY-) INCYTE PHARM INC.
 XX
 XX Wejher MG, Volkmuth W, Klingler TM, Lal P;
 XX WPI: 2000-558397/51.
 XX
 XX Novel gene associated with colon cancer and coexpressed with one or
 PT more known colon cancer genes in a number of biological samples, for
 XX use in gene therapy
 PS Claim 2; Page 30-31; 33pp; English.
 CC The present sequence represents a novel gene which is associated with
 CC colon disease. The gene is coexpressed with one or more colon cancer
 CC genes selected from carbonic anhydrase I, II and IV (CA I, II and IV),
 CC carcinoembryonic antigen family of proteins (cea), colorectal carcinoma
 CC tumour associated antigen (CO-029), down-regulated in adenoma (dra),
 CC fatty-acid binding protein (fabp), galectin (galect), glutathione
 CC peroxidase (gpx2), guanylin (guan), cyclokeratin 8 and 20 (ker 8 and 20),
 CC cadherin (cadher) or intestinal mucin (muc2). The present polynucleotide
 CC sequence is useful for diagnosing a disease or condition associated
 CC with altered expression of the gene. The polynucleotide and encoded
 CC polypeptides are useful for treating or preventing such diseases. They
 CC are useful for diagnosing colon cancer, metastatic colon cancer, atrophic
 CC gastritis, cholecystitis, Crohn's disease, irritable bowel syndrome and
 CC ulcerative colitis.
 XX
 SQ Sequence 2616 BP; 887 A; 488 C; 515 G; 726 T; 0 other;
 Query Match 62.1%; Score 135.4; DB 21; Length 2616;
 Best Local Similarity 76.5%; Pred. No. 1.3e-34; Mismatches 51; Indels 0; Gaps 0;
 Matches 166; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 1 GAACAATGGCCATCATGATGCTTTGGGGCCCTTTCATCAGGAAGATGAGCTGTCTCA 60
 DB 845 gaacaatggccatcatgctttggggccctttcacaggaagatgagctgtctca 914
 QY 61 GCGCTCCATCCAGCTTGAGAGTAAGGATTAACCTCCAGAACACCCAGTGAATGG 120
 DB 915 gcgctccatccagcttgagagtaaggattaacctccagaacacccagtgaatgg 974
 QY 121 CACAGTGAATCGGAGACAGCAGCGTGGGAAGAGACACTTGTCTTATCAGCTGGCAAC 180
 DB 975 ctacagtgaatcggagacagcagctgggaagagacacttgtcttattcagctggcaac 1034
 QY 181 GCAGCCTCCCAAAATCCTTCTCTGGAGATCCAGTGA 217
 DB 1035 gcagcctcccaaaatccttctctggagatccagtga 1071


```

RESULT 14
AAZ65095
ID AAZ65095 standard; cDNA; 3265 BP.
XX
AC AAZ65095;
XX
XX 05-APR-2000 (first entry)
XX
DE Membrane-bound protein FRO124 encoding cDNA.
XX
XX Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
XX pharmaceutical; receptor immunoadhesin; gene mapping; ss.
XX
OS Homo sapiens.
XX
XX MO9963088-A2.
XX
XX 09-DEC-1999.
XX
XX 02-JUN-1999; 99WO-US12252.
XX
XX 02-JUN-1998; 98US-0087607.
XX 02-JUN-1998; 98US-0087609.
XX 02-JUN-1998; 98US-0087759.
XX 03-JUN-1998; 98US-0087827.
XX 04-JUN-1998; 98US-0088021.
XX 04-JUN-1998; 98US-0088025.
XX 04-JUN-1998; 98US-0088028.
XX 04-JUN-1998; 98US-0088029.
XX 04-JUN-1998; 98US-0088030.
XX 04-JUN-1998; 98US-0088033.
XX 04-JUN-1998; 98US-0088326.
XX 05-JUN-1998; 98US-0088167.
XX 05-JUN-1998; 98US-0088202.
XX 05-JUN-1998; 98US-0088212.
XX 05-JUN-1998; 98US-0088217.
XX 09-JUN-1998; 98US-0088655.
XX 10-JUN-1998; 98US-0088722.
XX 10-JUN-1998; 98US-0088730.
XX 10-JUN-1998; 98US-0088734.
XX 10-JUN-1998; 98US-0088738.
XX 10-JUN-1998; 98US-0088740.
XX 10-JUN-1998; 98US-0088741.
XX 10-JUN-1998; 98US-0088742.
XX 10-JUN-1998; 98US-0088810.
XX 10-JUN-1998; 98US-0088811.
XX 10-JUN-1998; 98US-0088824.
XX 10-JUN-1998; 98US-0088825.
XX 10-JUN-1998; 98US-0088826.
XX 11-JUN-1998; 98US-0088858.
XX 11-JUN-1998; 98US-0088861.
XX 11-JUN-1998; 98US-0088863.
XX 11-JUN-1998; 98US-0088876.
XX 12-JUN-1998; 98US-0089090.
XX 12-JUN-1998; 98US-0089105.
XX 16-JUN-1998; 98US-0089440.
XX 16-JUN-1998; 98US-0089512.
XX 16-JUN-1998; 98US-0089514.
XX 17-JUN-1998; 98US-0089532.
XX 17-JUN-1998; 98US-0089538.
XX 17-JUN-1998; 98US-0089598.
XX 17-JUN-1998; 98US-0089599.
XX 17-JUN-1998; 98US-0089600.
XX 17-JUN-1998; 98US-0089653.
XX 18-JUN-1998; 98US-0089801.
XX 18-JUN-1998; 98US-0089907.
XX 18-JUN-1998; 98US-0089908.
XX 19-JUN-1998; 98US-0089947.
XX 19-JUN-1998; 98US-0089948.
XX 19-JUN-1998; 98US-0089952.
XX 22-JUN-1998; 98US-0090246.
XX 22-JUN-1998; 98US-0090252.

```

```

PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 25-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 12-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 19-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.

```


Tbu Apr 4 09:27:51 2002

Search completed: April 4, 2002, 03:17:53
Job time: 60225 sec

us-09-049-696-9.rng

This Page Blank (uspto)

AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/Genbank databases. Sunito Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing. Research Association for Biotechnology: cDNA library construction, 5' - 3' end one pass sequencing. Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES
source
 1. 2022
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COL02275"
 /clone_1id="COL"
 /tissue_type="colon"
 /note="cloning vector pME18SFL3"
 1. 2022
 /note="highly similar to AF127036 Homo sapiens calcium-activated chloride channel protein 1 (CACCI) mRNA"
BASE COUNT 612 a 472 c 453 g 485 t
ORIGIN
 Query Match 100.0%; Score 218; DB 9; Length 2022;
 Best local Similarity 100.0%; Pred. No. 3.3e-61;
 Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 GAACAATGGCCCTCATGATGCTTTGGGCGCCCTTCATCAGGAATGAGCTGTCTCA 60
 605 GAACATGGCCCTCATGATGCTTTGGGCGCCCTTCATCAGGAATGAGCTGTCTCA 664
 QY 61 GCGTCCATCCAGCTTGAGTAAGGATTAACCTCCAGAACAGCAGTGATGATG 120
 665 GCGTCCATCCAGCTTGAGTAAGGATTAACCTCCAGAACAGCAGTGATGATG 724
 Db 121 CACAGTATGCTGGACACACCGTGGAAGACACTTTGTTCTTATCAGCTGACAC 180
 725 CACAGTATGCTGGACACACCGTGGAAGACACTTTGTTCTTATCAGCTGACAC 784
 QY 181 GCAGCTCCCAAAATCCTCTCTGGGATCCAGTGGAC 218
 785 GCAGCTCCCAAAATCCTCTCTGGGATCCAGTGGAC 822
RESULT 2
LOCUS AF127036 2826 bp mRNA PRI 10-AUG-1999
DEFINITION Homo sapiens calcium-activated chloride channel protein 1 (CACCI)
ACCESSION AF127036
VERSION AF127036.1 GI:4585468
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Agnel, M., Vermet, T. and Culouscou, J. M.
TITLE Identification of three novel members of the calcium-dependent chloride channel (CACCI) family predominantly expressed in the digestive tract and trachea
JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792
REFERENCE 2 (bases 1 to 2826)

AUTHORS Agnel, M. and Culouscou, J. M.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrieres, Rueil-Malmaison 92500, France
FEATURES
source
 Location/Qualifiers
 1. 2826
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="small intestine; colon"
 1. 2826
 /gene="CACCI"
 5. 2749
 /gene="CACCI"
 /note="bovine epithelial chloride channel homolog"
 /codon_start=1
 /product="calcium-activated chloride channel protein 1"
 /protein_id="A025487.1"
 /db_xref="GI:4585469"
 /translation="MGPFKSSVFLLIHLELGAISNLLIOLNNGYEGIVVAIDPNP
 EDELIIQIKDMVTOASLYLEFATGEFFKRNVAIIPETWKTADYVRKLEYKNA
 DVIVAEETPPGNDPEPTTEOMGNGEKEGRILIPDPFAGKRIAEYGGRAHYEMATL
 LRMGVPEYNNDEKFEYLSNGRIQAVRCSAGITGNVYKKGCGSCYTRKCFNKVGL
 YKGCERFVLSROTERASTIMPAQVDSIVERCTEONNNKAPKONKCNIRSWPT
 RDSDEFKTTPTTPPPPTFSLQIQRTVCLVLDKSGSMATGNRLRNQAGLEFL
 LQVLELSWGMVTFDSAAHYOSELQIINGSDRDLAKRIPAAASGTSICGLRBA
 FTVIRKKYPTDSEIVLITDGEDNTISGCENYVSGAIIHTVALGSAOELSELK
 MTGGLQVYASDOVONNGILDAFGLSSGNCASVRSRISLSEGLITLQNSOMNGTVY
 DSYGKDTLFLITWTQPPQITLMDPSGOKGCPYVKNKRMAYIOTPIAKCTWY
 SIQASSTDLITLITVSRASNAITLPIITVTSKTKNDTSKFPSPVLYVYAIRGASPTLKA
 SVAILLESVNGKVTLELNDGAGADATKDDGVSRFTYDINGRISVVRALGVNA
 AARRRVIPQSGALYIPGWIENDEIOMNPPRPEIKMDVOHKOVCESRTSGSFVS
 DVNPAPIPDLFPQITDLKAEIHGSLINLTWAPDDYDHGAHYITRISLID
 LRDFNLSLVNTTALIPKANSSEVLEFPENTFENGDLFLAIOAVKVDLKSFI
 SNLARVSLFIPQPPPTPPSPDETSAFCPNHINSTIPGIHLIKIMKMGELQSLIA"
BASE COUNT 875 a 623 c 632 g 696 t
ORIGIN
 Query Match 100.0%; Score 218; DB 9; Length 2826;
 Best local Similarity 100.0%; Pred. No. 3.4e-61;
 Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 GAACAATGGCCCTCATGATGCTTTGGGCGCCCTTCATCAGGAATGAGCTGTCTCA 60
 1396 GAACATGGCCCTCATGATGCTTTGGGCGCCCTTCATCAGGAATGAGCTGTCTCA 1455
 QY 61 GCGTCCATCCAGCTTGAGTAAGGATTAACCTCCAGAACAGCAGTGATGATG 120
 1456 GCGTCCATCCAGCTTGAGTAAGGATTAACCTCCAGAACAGCAGTGATGATG 1515
 Db 121 CACAGTATGCTGGACACACCGTGGAAGACACTTTGTTCTTATCAGCTGACAC 180
 1516 CACAGTATGCTGGACACACCGTGGAAGACACTTTGTTCTTATCAGCTGACAC 1575
 QY 131 GCAGCTCCCAAAATCCTCTCTGGGATCCAGTGGAC 218
 1576 GCAGCTCCCAAAATCCTCTCTGGGATCCAGTGGAC 1613
RESULT 3
LOCUS AX193489 3311 bp DNA PAT 15-AUG-2001
DEFINITION Sequence 1056 from Patent W00149716.
ACCESSION AX193489
VERSION AX193489.1 GI:15211440
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Xu, J., Lodes, M. J., Secretist, H., Benson, D. R., Meagher, M. J.,

TITLE
Stolk, J.A., King, G.E., Wang, T. and Jiang, Y.
Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
Patent: WO 0149716-A 1056 12-JUL-2001;
JOURNAL
CORIXA CORPORATION (US)

FEATURES
Source
Location/Qualifiers
1. .3311
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 100.0%; Score 218; DB 6; Length 3311;
Best Local Similarity 100.0%; Pred. No. 3.5e-61;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAAATGGCCATGATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGCTGTCTCTCA 60
DB 1743 GACAAATGGCCATGATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGCTGTCTCTCA 1802
QY 61 GCGCTCCATCCAGCTTGAAGGATTAACCCCTCAGAACAGCCAGTGGATGATG 120
DB 1803 GCGCTCCATCCAGCTTGAAGGATTAACCCCTCAGAACAGCCAGTGGATGATG 1862
QY 121 CACAGTATCGTGGACAGCAGCCGTGGAAAGACACTTTGTTCTTATCAGCTGACAC 180
DB 1863 CACAGTATCGTGGACAGCAGCCGTGGAAAGACACTTTGTTCTTATCAGCTGACAC 1922
QY 181 GCAGCCTCCCAAAATCCTTCTCTGGGATCCAGTGAGC 218
DB 1923 GCAGCCTCCCAAAATCCTTCTCTGGGATCCAGTGAGC 1960

RESULT 4

LOCUS AF039400 3311 bp mRNA PRI 14-DEC-1998
DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLIC1) mRNA,
complete cds.
ACCESSION AF039400
VERSION AF039400.1 GI:4009457
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 3311)
Gruber, A.D., Elble, R.C., Ji, H.L., Schreier, K.D., Fuller, C.M. and
Pauli, B.U.
TITLE Genomic cloning, molecular characterization, and functional
analysis of human CLIC1, the first human member of the family of
Ca²⁺-activated Cl⁻ channel proteins
JOURNAL Genomics 54 (2), 200-214 (1998)

MEDLINE
REFERENCE 2 (bases 1 to 3311)
Gruber, A.D., Elble, R. and Pauli, B.U.

AUTHORS Direct Submission
TITLE Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
Cornell University, Ithaca, NY 14853, USA

FEATURES
Source
Location/Qualifiers
1. .3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/tissue_type="small intestine"
1. .3311
/gene="hCLIC1"
352. .3096
/gene="hCLIC1"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"

gene
CDS

1. .3311
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
/tissue_type="small intestine"
1. .3311
/gene="hCLIC1"
352. .3096
/gene="hCLIC1"
/note="transmembrane glycoprotein"
/codon_start=1
/product="calcium-dependent chloride channel-1"

/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPEKSSVFLILHLLEGALSNSLIQLNNNGYEGIVADPNVP
EDFTLLOOIKQWQASLYLPMATGKRKFQFVAVLIPEWTAKTKADYVAPKLETVKA
DYLVASPEKPNDEPYTRDMCGCKEGERIHLPPIAGKALAEVGPCKAVHEMAH
LRMGVDEYNDEKFTLSNGRIQAVRCSVAGITGVNKKCGGSCYTRKTNKVTGL
YEKGGEFVQSRQTERKASIMFAQHDVSIYFCTEQHNHNEAPNOKNOKNLSRTEVI
RDSDEKFTTPTTQPPNPTFSLDIGITVCLVDKSGSNATGRNLNNOGOLF
LQTVELGSGWGVTFDSAAHVOSELIQINSQSDRPLAKRLPAAASGCTSGSRSRA
FVIRKKYPTQSGSEIILLDGEDNTISGFENVKQSGAIIHVVALGSAOBELELSK
MTGQIQTASDVOYNNGLIDAGCALSSNGCAVSQSIQLESGLTLNOSNNNGYIV
DSTGKDTLFLITWTQPPQIILMPDSGQGGEVVDKNTKAVLIQIPEIAVGTWKY
SLQASSQTLTLTVTSRASNATLPITVTSKTRKDSKEPSLVVANNROGASPLRA
SVTALIESYNGKTVTLELDNGADATDQDVSRFTYTDNRYSVYKVALGAVN
AARVVITPOOSALYIPGWIENDEIOMNPREDIKDVOYHOKVCFSPRSGGSPVAS
DYPNAPIDPLPPGOITPLDKARIGHGSLINLTWTAQGDYDHGTAHKYIIRISLID
LBDKRNESIQVNTPTLIPREANSEVPLFKPNITPENCTDLFLAQAVDKYDKLSEI
SNIAVSLFIPQTPETPSDETAPCPNHIINSTIRPQIHILKIMMKWIGLOLSIA

BASE COUNT 1028 a 692 c 742 g 849 t
ORIGIN

Query Match 100.0%; Score 218; DB 9; Length 3311;
Best Local Similarity 100.0%; Pred. No. 3.5e-61;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAAATGGCCATGATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGCTGTCTCTCA 60
DB 1743 GACAAATGGCCATGATGATGCTTTGGGCGCCCTTTCATCAGAAATGAGCTGTCTCTCA 1802
QY 61 GCGCTCCATCCAGCTTGAAGGATTAACCCCTCAGAACAGCCAGTGGATGATG 120
DB 1803 GCGCTCCATCCAGCTTGAAGGATTAACCCCTCAGAACAGCCAGTGGATGATG 1862
QY 121 CACAGTATCGTGGACAGCAGCCGTGGAAAGACACTTTGTTCTTATCAGCTGACAC 180
DB 1863 CACAGTATCGTGGACAGCAGCCGTGGAAAGACACTTTGTTCTTATCAGCTGACAC 1922
QY 181 GCAGCCTCCCAAAATCCTTCTCTGGGATCCAGTGAGC 218
DB 1923 GCAGCCTCCCAAAATCCTTCTCTGGGATCCAGTGAGC 1960

RESULT 5

LOCUS AB017156 2937 bp mRNA ROD 10-NOV-1999
DEFINITION Mus musculus gob-5 mRNA, complete cds.
ACCESSION AB017156
VERSION AB017156.1 GI:3721911
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (sites)
Komiyama, T., Tanigawa, Y. and Hirohashi, S.
TITLE Cloning and identification of the gene gob-5, which is expressed in
intestinal goblet cells in mice
JOURNAL Biochem. Biophys. Res. Commun. 255 (2), 347-351 (1999)

MEDLINE
REFERENCE 2 (bases 1 to 2937)
Komiyama, T., Tanigawa, Y. and Hirohashi, S.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1998) to the DDBJ/EMBL/GenBank databases, Tohru
Komiyama, ERATO, JST, Genome Asymmetry Group, Doi Biosymmetry
Project, Tokodai 5-9-9, Tsukuba, Ibaragi 300-2635, Japan
(E-mail:tkom@bio.eriato.tlc-net.co.jp, Tel:81-298-48-1515, Fax:81
298-47-8901)

FEATURES
Source
Location/Qualifiers
1. .2937
/organism="Mus musculus"
/db_xref="taxon:10090"

Query Match	67.98;	Score 148;	DB 9;	Length 35278;
Best Local Similarity	100.0%;	Pred. No. 6.7e-38;		
Matches 148;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least Q40
Consensus quality: 162610 bases at least Q30
Consensus quality: 163328 bases at least Q20
Insert size: 163991; sum-of-connigs
Insert size: 163373; 6.6% error; agarose-fp
Quality coverage: 5.19x in Q20 bases; sum-of-connigs
Quality coverage: 5.21x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 34236: contig of 34236 bp in length
* 34237 34336: gap of 100 bp
* 34337 47372: contig of 13036 bp in length
* 47373 47472: gap of 100 bp
* 47473 67025: contig of 19553 bp in length
* 67026 67125: gap of 100 bp
* 67126 71239: contig of 4114 bp in length
* 71240 71339: gap of 100 bp
* 71340 83903: contig of 12564 bp in length
* 83904 84003: gap of 100 bp
* 84004 91357: contig of 7354 bp in length
* 91358 91457: gap of 100 bp
* 91458 96452: contig of 4995 bp in length
* 96453 96552: gap of 100 bp
* 96553 148189: contig of 51637 bp in length
* 148190 148289: gap of 100 bp
* 148290 159778: contig of 11489 bp in length
* 159779 159878: gap of 100 bp
* 159879 164981: contig of 5013 bp in length.
*
Location/Qualifiers
1.164631
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```

/notice="assembly_fragment:0002"
fragment_chain:1
clone_end:17
vector_side:left"
34337.47372
/notice="assembly_fragment:0031"
fragment_chain:1"
47473.67025
/notice="assembly_fragment:0192"
fragment_chain:1"
67126.71239
/notice="assembly_fragment:0209"
fragment_chain:1"
71340.83903
/notice="assembly_fragment:0221"
fragment_chain:1"
84004.91357
/notice="assembly_fragment:0129"
fragment_chain:1"
91458.96452
/notice="assembly_fragment:0139"
fragment_chain:1"
96553.148189
/notice="assembly_fragment:0204"
fragment_chain:1"
148290.159778
/notice="assembly_fragment:0042"
fragment_chain:1"
159879.164891

```


/note="assembly_fragment:00773
fragment_chain:1
clone_end:SP6
vector_side:right"
BASE COUNT 48381 a 32179 c 32357 g 51067 t 907 others
ORIGIN

Query Match 67.9%; Score 148; DB 2; Length 164891;
Best Local Similarity 100.0%; Pred. No. 8e-38;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CAGCTTGAGAGTAAGGATTAACCTCCAGAACACAGCCAGTGTGATGATGATGATC 130
|||||
DB 55363 CAGCTTGAGAGTAAGGATTAACCTCCAGAACACAGCCAGTGTGATGATGATGATC 55304
|||||
QY 131 GTGACACACCGGTGGGAAAGACACTTTGTTCTTATCAGCTGACCAACGAGCTCC 190
|||||
DB 55303 GTGACACACCGGTGGGAAAGACACTTTGTTCTTATCAGCTGACCAACGAGCTCC 55244
|||||
QY 191 CAAATCTTCTCTGGGATCCAGTGAC 218
|||||
DB 55243 CAAATCTTCTCTGGGATCCAGTGAC 55216
|||||

RESULT 8
HSJ651E10 113764 bp DNA PRI 27-MAY-2000
LOCUS Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1.
DEFINITION complete sequence.
ACCESSION AL122002
VERSION AL122002.16 GI:8247274
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 113764)
AUTHORS Brown,A.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
Request: clonerequest@sanger.ac.uk
On Jun 4, 2000 this sequence version replaced gi:8247030.
During sequence assembly data is compared from overlapping clones.
When differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; SW, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1

RP4-651E10 is from the library RPCI-4 constructed at the Roswell
Park Cancer Institute by the group of Peter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pcypac2

This sequence is the entire insert of clone RP4-651E10.

FEATURES
SOURCE location/Qualifiers
1..113764
/organism="Homo sapiens"

/db_xref="taxon:9606"
/chromosome="1"
/map="p22.3-31.1"
/clone="RP4-651E10"
/clone_11b="RPCI-4"
BASE COUNT 34890 a 21985 c 22351 g 34534 t
ORIGIN

Query Match 67.2%; Score 146.4; DB 9; Length 113764;
Best Local Similarity 99.3%; Pred. No. 2.6e-37;
Matches 147; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 71 CAGCTTGAGAGTAAGGATTAACCTCCAGAACACAGCCAGTGTGATGATGATGATC 130
|||||
DB 26351 CAGCTTGAGAGTAAGGATTAACCTCCAGAACACAGCCAGTGTGATGATGATGATC 26410
|||||
QY 131 GTGACACACCGGTGGGAAAGACACTTTGTTCTTATCAGCTGACCAACGAGCTCC 190
|||||
DB 26411 GTGACACACCGGTGGGAAAGACACTTTGTTCTTATCAGCTGACCAACGAGCTCC 26470
|||||
QY 191 CAAATCTTCTCTGGGATCCAGTGAC 218
|||||
DB 26471 CAAATCTTCTCTGGGATCCAGTGAC 26498
|||||

RESULT 9
AF095584 3079 bp mRNA MAM 13-OCT-2000
LOCUS Sus scrofa epithelial chloride channel protein (AECC) mRNA,
DEFINITION complete cds.
ACCESSION AF095584
VERSION AF095584.1 GI:6002645
KEYWORDS
SOURCE pig.

ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
REFERENCE 1 (bases 1 to 3079)
AUTHORS Gaspar,K.J., Racette,K.J., Gordon,J.R., Loewen,M.E. and
Forsyth,G.W.

TITLE Cloning a chloride conductance mediator from the apical membrane of
porcine ileal enterocytes
JOURNAL Physiol. Genomics (Online) 3 (2), 101-111 (2000)
MEDLINE 20473747
PUBMED 11015605

REFERENCE 2 (bases 1 to 3079)
AUTHORS Gaspar,K.J., Gabriel,S.E., Racette,K.J. and Forsyth,G.W.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-1998) Veterinary Physiological Sciences,
University of Saskatchewan, 52 Campus Drive, Saskatoon, SK S7N 5B4,
Canada

FEATURES
SOURCE location/Qualifiers
1..3079
/organism="Sus scrofa"
/db_xref="taxon:9823"
/tissue_type="ileal mucosa"
1..3079
/gene="AECC"
133..2886
/gene="AECC"
/function="involved in exocrine chloride conductance"
/codon_start=1
/product="epithelial chloride channel protein"
/protein_id="AAP00077.1"
/db_xref="GI:6002645"

/translation="MGSFRRSLFLVLHLLEGASNSLIQNGNGYGVIAIDPNP
EDERLIONIKDMWKASPYLFETAEKRFYFNVAIILPAKWKAPYVKKRLETKNA
DVVYTEPNPENDPYTEOMNGCEKSEKIYFPDPAKKVLOYGOGRFVEMAH
LRMGVFNENNEOKFVLSNKKKEPVISAAIRGNVLPDGGSCVTRPCPADRVGL
FOKRCERIPDQSEKASIMPAOSIDPYVECKRKNKRNPNQONKCNIRSWYI
QDSERKTPPTPTTPPTSLQIQIRIVCLVDSGSGTGGRLKRLNQAKLFL
LOTVEGAWGMVAFDSAAAYVSELVQINSAAERDALRLPTAASGTSISGLRSA

FEATURES	location/qualifiers
source	1..1895
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="COL06357"

```
source 1. 3204
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /tissue_type="colon"
gene 1. 3204
```


CDS

/gene="CaCC2"
29..2782
/gene="CaCC2"
/note="bovine epithelial chloride channel homolog"
/product="calcium-activated chloride channel protein 2"
/protein_id="A48398.1"
/db_xref="GI:5726289"
/translation="MGFRGFVFLVLLHQSNTSFILKNNGEDIVIVDPSPVE
DEKIEQEDVATTSYLFTEATEKREFFKNVSLIPENKKNPPKPKHKNHHA
VIVAPPLTGRDEPTKOTEGEKEVILHTPOLTKKKNPPKPKHKNHHA
RMGVDEYNEDQPTFRASKIKIATRCASGSRNRYICGGSLASCRIDSTTK
YKDCQFFPDVQTEKASIMQSDISVVERCNEHTHQAASLQNICNRSSTEV
SNSDFKNTIPMVPPTPPVSLKIRORIVCLVDKSGMGKDRNMQAHL
LOTYNGSMGVMHFDSTATVNLKIQKSSDERNTLAAGLPYPLGSGISGKYA
FOVIGELHSDSEVLLTDEGENTASSIDEVQSGAIVHFIAGRADEAVIEMS
KITGSHFYVDEQNGNLIDAFALTSQNDLSQKSLKGLTNSNMMDTVI
IDSTYKDTFFELITNSLPSTISIMDESGTITMNTVDATSKMALISIPGAKVTA
YNLOAKNPELTITVTSRAANSVPPTVNAKNNKDVNSPSPMIVAEILQGVV
LGANTVAFIESONGTEVLELDNGADSEFNKQVSRFYATYENGRYSILKRVAG
GANTARLRLRPPLNAAIIPGVVNGEIEANPREIDEDQTLDEDSRAGGAV
VSOVSLPDPPOYPSOITDADATVHEKIIITWTAPGDNFDGKVOYRITRISATL
DLRSPDALQVNTDLSKRNKSESAFENISENATHIRIATISIKSITK
VSNIAQVTLFIQANPDIDPTPTPTPKSHSGVNIITLVLSVIVIVNFTLS
TTI"

BASE COUNT 1098 a 594 c 633 g 879 t

ORIGIN

Query Match 62.1%; Score 135.4; DB 9; Length 3204;
Best Local Similarity 76.5%; Pred. No. 7.6e-34;
Matches 166; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

OY 1 GAACATGCGCTCATGATGCTTTTGGGCGCTTCATCAGGAAGATGAGCTGCTCA 60
DB 1423 GAACATGCGCTCATGATGCTTTTGGGCGCTTCATCAGGAAGATGAGCTGCTCA 1482
OY 61 GCGCTTCATCAGCTTGAGAGTAAAGGATTAACCTCCAGAACCCAGCTGATGATG 120
DB 1483 GAAGTCCTTCAGCTGAGAAAGGATTAACCTGATTAATGCTGATGATGATG 1542
OY 121 CACAGTGCAGCGAGACAGCGTGGGAAAGACACTTGTCTTATCAGCTGAGCAAC 180
DB 1543 CACTGTCAATTAATGATGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1602
OY 181 GCAGCTCCGCCCAATCTCTCTGGGATCCAGTGA 217
DB 1603 TCTGCTCCAGTATTTCTCTGGGATCCAGTGA 1639

RESULT 12
LOCUS AK000072 3221 bp mRNA PRI 22-FEB-2000
DEFINITION Homo sapiens cDNA FLJ20065 fis, clone COL01613, highly similar to
ECLC BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.
ACCESSION AK000072
VERSION AK000072.1 GI:7019922
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Ohashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T., and Sugano, S.
TITLE Unpublished (2000)
JOURNAL 2 (bases 1 to 3221)
REFERENCE Sugano, S., Suzuki, Y., Ota, T., Ohashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T., and Nakamura, Y.
AUTHORS Direct Submission
TITLE Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio

Sugano, Institute of Medical Science, University of Tokyo, Department
of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: chn@leims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)
COMMENT MEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5' & 3' end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
source
1..3221
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01613"
/clone_lib="COL"
/tissue_type="colon"
/note="cloning vector pME18SFL3"
misc_feature
1..3221
/note="highly similar to ECLC_BOVIN EPITHELIAL CHLORIDE
CHANNEL PROTEIN"

BASE COUNT 1105 a 600 c 634 g 882 t

ORIGIN

Query Match 62.1%; Score 135.4; DB 9; Length 3221;
Best Local Similarity 76.5%; Pred. No. 7.6e-34;
Matches 166; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

OY 1 GAACATGCGCTCATGATGCTTTTGGGCGCTTCATCAGGAAGATGAGCTGCTCA 60
DB 1423 GAACATGCGCTCATGATGCTTTTGGGCGCTTCATCAGGAAGATGAGCTGCTCA 1496
OY 61 GCGCTTCATCAGCTTGAGAGTAAAGGATTAACCTCCAGAACCCAGCTGATGATG 120
DB 1437 GAAGTCCTTCAGCTGAGAAAGGATTAACCTGATTAATGCTGATGATGATG 1556
OY 121 CACAGTGCAGCGAGACAGCGTGGGAAAGACACTTGTCTTATCAGCTGAGCAAC 180
DB 1543 CACTGTCAATTAATGATGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1616
OY 181 GCAGCTCCGCCCAATCTCTCTGGGATCCAGTGA 217
DB 1617 TCTGCTCCAGTATTTCTCTGGGATCCAGTGA 1653

RESULT 13
LOCUS AX092338 3265 bp DNA PAT 21-MAR-2001
DEFINITION Sequence 69 from Patent WO0116318.
ACCESSION AX092338
VERSION AX092338.1 GI:13444483
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Eaton, D.L., Filvaroff, E., Gerritsen, M.E., Goddard, A.,
Goddard, C.J., Grimaldi, C.J., Gurney, A.L., Watanabe, C.K., and
Wood, W.I.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0116318-A 69 08-MAR-2001;
Genentech, Inc. (US)
FEATURES
source
1..3265
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1159 a 596 c 632 g 878 t

CGSFYVSGARPGNHPSVLPFNKIIDLFAKFKDHIQLSWTAFLANLIDKANSYITIR
SKSPFLDLQKDFDNATLWNTSSILKPKAGSGSDNFEEKPEPRLIENGTFYIAVOAINEE

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:27:12 ; Search time 16681 Seconds
(without alignments)
116.599 Million cell updates/sec

Title: US-09-049-696-4
Perfect score: 181
Sequence: 1 CAAGAATGCGCATTCATTA.....ACAAGCAAAATCAAAATGC 181

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 segs, 537289281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: em_estfun:*
- 2: em_esthum:*
- 3: em_estlin:*
- 4: em_estom:*
- 5: em_estpl:*
- 6: em_estda:*
- 7: em_estro:*
- 8: em_estov:*
- 9: em_hic:*
- 10: qb_est1:*
- 11: qb_est2:*
- 12: qb_hic:*
- 13: qb_gss:*
- 14: em_gss_fun:*
- 15: em_gss_hum:*
- 16: em_gss_inv:*
- 17: em_gss_pln:*
- 18: em_gss_pro:*
- 19: em_gss_rnd:*
- 20: em_gss_vit:*
- 21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	177.8	98.2	480	10	AA581198 nd38c07.r
2	167	92.3	460	10	AA750670 CM2-CN003
3	94.8	52.4	2915	12	AK007466 Mus muscu
4	94.8	52.4	2933	12	AK008659 Mus muscu
5	88.6	49.0	307	10	AK008659 Mus muscu
6	87.6	48.4	900	11	BF578833 CM0-CN000
7	71.2	39.3	427	10	AA498535 vha40f12.r
8	71.2	39.3	640	11	BF320704 u255f02.y
9	68	37.6	547	10	BF380745 601270855
10	66.4	36.7	935	11	BF099600 601751621
11	66.4	36.7	949	11	BF099712 601751347
12	59.2	32.7	972	11	BG173173 602335553

13	58.8	32.5	602	10	AL602490 DKFZp686P
14	58.8	32.5	758	10	AL598508 DKFZp313M
15	57.2	31.6	580	11	BI342391 370140 MA
16	54.4	30.1	858	11	BI110624 602896268
17	54.2	29.9	806	11	BG562080 602826824
18	53	29.3	1047	11	BG247065 602359723
19	49.4	27.3	823	11	BG741082 602631626
20	45.8	25.3	395	13	A2465176 1M0274523
21	45	24.9	305	10	B2440298 BB240298
22	42.4	23.4	402	10	BE656570 UT-M-BHO-
23	41	22.7	650	13	BH035021 RPT1-24-3
24	39.4	21.8	744	13	A2855898 2M0160E12
25	38.2	21.1	545	11	BF969353 602271617
26	37.6	20.8	429	11	BG286687 602382666
27	37.6	20.8	657	11	BE275492 601121331
28	37.2	20.6	608	13	A2897106 RPT1-24-1
29	37	20.4	174	11	BE636885 NF073C07L
30	36.8	20.3	184	11	BI289861 UT-R-DKO-
31	36.8	20.3	457	13	A2152659 SP_0034_B
32	36.8	20.3	471	13	A2152460 SP_0034_B
33	36.6	20.2	750	11	BG261362 602373135
34	36.4	20.1	874	11	BG291548 602385623
35	36.2	20.0	254	10	BE148660 MRO-H1024
36	36	19.9	205	10	AA290063 VB50a10.r
37	36	19.9	507	11	BG966381 602832873
38	36	19.9	511	11	BF794844 602256702
39	36	19.9	849	13	A2189337 SP_1014_A
40	36	19.9	901	16	BF237690 601841876
41	35.8	19.8	309	10	AI351964 qd23a12.x
42	35.8	19.8	374	11	BI055590 CM3-GN046
43	35.6	19.7	764	13	AQ254324 CpG0766A
44	35.6	19.7	355	13	AQ314468 RPT11-10
45	35.6	19.7	435	13	AO776782 HS_5570_A

ALIGNMENTS

RESULT 1
LOCUS AA581198 480 bp mRNA EST 05-JAN-1998
DEFINITION nd38c07.r1 NCI-CGAP.C01 Homo sapiens cDNA clone IMAGE:802572.5,
similar to FR:G1184066 G1184066 CALCIUM-ACTIVATED CHLORIDE CHANNEL.
// mRNA sequence.
ACCESSION AA581198
VERSION AA581198.1 GI:23586970
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 480)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies Inc., David Krizman,
Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNLT at:
www.bio.lnlnl.gov/bdrp/image/image.html
Seq primer: -28m13 rev1 ET from Amerisham
High quality sequence stop: 405.
Location/Qualifiers
1.480
/organism="Homo sapiens"
/db_xref="taxon:9606"


```

/clone="IMAGE:802572"
/clone_1b="NCI-CCAP_Co1"
/tissue="type="bulk tumor"
/lab_host="DH10B"
/site="Organ: colon; Vector: PCMV-SPORT2; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
168 a 101 c 97 g 114 t
BASE COUNT
ORIGIN

```

Query Match	98.2%;	Score 177.8;	DB 10;	Length 480;
Best Local Similarity	98.9%;	Pred. No. 2.5e-29;		
Matches 179;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY      I CAAAGATGCACATTCAATAAAGTAACGAGCCTTATGA AAAAGCATGTGAGTTTGTCCT   60
        |||
Db    102 CAAAAGATGCACATTCAATAAAGTACAGGACTTATGAAAAAGCATGTGAGTTTGTCCT   161
```

09 CCAATCCCGCAGACGAGAGGCTTATATATGTTGCACMACATGTGATTCTAAGT 120
|||
Db 162 CCAATCCCCGCCAGAAGAGCAAGGCTTCATTAATGTTGCACACATCTTGATTCATAAGT 221

0y 121 TGAATTCGTACAGACAAAACCACAAAGAAGCTCCAAACAAGCAAAATCATAAATG 180
|||
Db 222 TGAATTCTGTACAGAACAAACCAACAAGAAGCTCCAACAAGCAAAATCATAAATG 281

QY	181	C	181
Db	282	C	282

RESULT	2
AM750670/c	
LOCUS	460 bp mRNA
DEFINITION	28-APR-2000
ACCESSION	CM2-CM0039-110100-069-c01 CM0039 Homo sapiens CDNA, mRNA sequence
AM750670	

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 460)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
URL <http://www.ncbi.nlm.nih.gov/PMCID/PMC100065/>

COMMENT

Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpcos@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl1=CM2&st2=CM2-CN0039-101010-069-c01&t3=2000-01-11&t4=1>)
Seq primer: puc-18 forward

FEATURES	Location/Qualifiers
source	1. .460

1.460

```
/organism="Homo sapiens"
```

```
/ab_xrel="Laxon:9606"
/cjone_1:b="CN0039"
```

```
/dev stage="Adult"
```

```

/acc=acc      name=
/notes="Organ: colon normal: Vector: puc18: Site 1: Smat:

```

Site-2: Smal; A mini-library was made by cloning products derived from ORSRES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions.

BASE COUNT 114 a 96 c 100 g 150 t
ORIGIN

Query Match	92.3%	Score 167	DB 10	Length 460
Best local similarity	100.0%	Pred. No. 5.5e-2		
Matches 167	Conservative 0	Mismatches 0	Indels 0	Gaps 0

15 TCATTAAGATACAGAGCTCTATGAAAAAGATGTGAGTTGTTCTCCAATCCCGCAGA 74
 30 |||||
 45 TCATTAAGATACAGAGCTCTATGAAAAAGATGTGAGTTGTTCTCCAATCCCGCAGA 401

QY 75 CGGAGAAGGCTTCTATATGTTTGCACACATGTTGATCTATAGTTGAATTCGTGCAC 134
 |||||
 Db 400 CGGAGAAGGCTTCTATATGTTTGCACACATGTTGATCTATAGTTGAATTCGTGCAC 341

QY 135 AACAAATCCACACAAAGAGCTCCAACAACAGCAAAATCAAAATGC 181
 Db 340 AACAAATCCACACAAAGAGCTCCAACAACAGCAAAATCAAAATGC 294

RESULT : 3
AK007465
LOCUS :

Accession	Definition	LOCUS	2915 bp	mRNA	HTC	05-JUL-2001
AK007465	Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810012P03, full insert sequence.	AK007466				

MODIFIED: 2010-07-10
 VERSION: AK007466.1 GI:12841032
 KEYWORDS: CAP trapper.
 SOURCE: Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to
 mRNA cloned 1b, seven full length cDNAs cloned from the

ORGANISM

REFERENCE
1 (bases 1 to 2915)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA

REFERENCE

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

MEDLINE

REFERENCE	3 (bases 1 to 2915)
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komori, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M.,

Samino, T., Ishii, Y., Nakamura, S., Hazama, M., Nishitane, T., Harada, A., Yamamoto, R., Matsunoto, H., Sekiguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Ozawa, Y., Ienaga, M., Ohara, E., Watabiki, K., Yoneda, Y., Ishikawa, T., Tazawa, K., Tanaka, T., Matsura, S., Kawai, T., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system—334-format sequencing pipeline with 384 multichannelillary sequencer Genome research, 10 (11), 1757-1771 (2000)

MEDLINE 20530913
PUBMED 11076861

REFERENCE 4 (base

AUTHORS The RIKE
DAVIDSON

FANTOM C
TITTE
Euro+300

Functional annotation of a full-length mouse cDNA collection

RESULT	5
AM577823	
LOCUS	AM577823
DEFINITION	CW0-CN0004-130100-139-h03 CN0004 Homo sapiens CDNA, mRNA sequence
ACCESSION	AM577823
VERSION	AM577823.1 GI:7252872
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 203 a 154 c 138 g 142 t 3 others

Query Match 39.3%; Score 71.2; DB 11; Length 640;
Best Local Similarity 66.0%; Pred. No. 3.3e-06;
Matches 103; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 26 ACAGAGCTCTATGAAAAAGATGTGATTTCTTCACATCCCGCAGAGAGAGCT 85
DB 89 ACACGGCTGTATGAACCAACCAATGTATCCGACAAATACACGCGTGGGCC 148

QY 86 TCTATATGTTGCACACATGTGATCTATGTTGATTCGTGACAGACAAACAC 145
DB 149 TCCTAATGTTGATGCAAAACCTTATTTCTGTGTAATTTTGCACAGAAAAACAC 208

QY 146 AACAAAGAGCTCCAAACAGCAAAATCAAAATGC 181
DB 209 AATGACAGAGCCCAACCTACAAACAAATGTGC 244

RESULT 9
BE380745 547 bp mRNA EST 21-JUL-2000
LOCUS 601270855F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3600205 5',
DEFINITION mRNA sequence.
ACCESSION BE380745.1 GI:9326110
VERSION BE380745.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 547)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM8783 row: n column: 14
High quality sequence stop: 544.

FEATURES
source location/Qualifiers

1. 547

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3600205"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 160 a 135 c 130 g 121 t 1 others

Query Match 37.6%; Score 68; DB 10; Length 547;
Best Local Similarity 64.7%; Pred. No. 1.7e-05;
Matches 101; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 26 ACAGAGCTCTATGAAAAAGATGTGATTTCTTCACATCCCGCAGAGAGAGCT 85
DB 247 ACACGGCTGTATGAACCAACCAATGTATCCGACAAATACACGCGTGGGCC 306
QY 86 TCTATATGTTGCACACATGTGATCTATGTTGATTCGTGACAGACAAACAC 145
DB 337 TCCTAATGTTGATGCAAAACCTTATTTCTGTGTAATTTTGCACAGAAAAACAC 366

QY 146 AACAAAGAGCTCCAAACAGCAAAATCAAAATGC 181
DB 357 AATGACAGAGCCCAACCTACAAACAAATGTGC 402

RESULT 10
BF099600 935 bp mRNA EST 19-OCT-2000
LOCUS 601751621F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3979468 5',
DEFINITION mRNA sequence.
ACCESSION BF099600.1 GI:10882126
VERSION BF099600.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 935)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM9173 row: 1 column: 05
High quality sequence stop: 595.

FEATURES
source location/Qualifiers

1. 935

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3979468"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 292 a 239 c 226 g 178 t

Query Match 36.7%; Score 66.4; DB 11; Length 935;
Best Local Similarity 64.1%; Pred. No. 3.3e-05;
Matches 100; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 26 ACAGAGCTCTATGAAAAAGATGTGATTTCTTCACATCCCGCAGAGAGAGCT 85
DB 244 ACACGGCTGTATGAACCAACCAATGTATCCGACAAATACACGCGTGGGCC 343
QY 86 TCTATATGTTGCACACATGTGATCTATGTTGATTCGTGACAGACAAACAC 145

DB 344 TCCATATATTTATGCAAGGCAATCTGCTGCAATTTTGACAGAAATAACAC 403
OY 146 AACAAAGAGCTCAACAAAGCAAAATCAAAATGC 181
DB 404 AATGCAGAAAGCCCAAACTACAAAACAAATGTGC 439

RESULT 11
BF099712 949 bp mRNA EST 19-OCT-2000
LOCUS 601751347F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3979187 5',
DEFINITION mRNA sequence.
ACCESSION BF099712 GI:10882238
VERSION BF099712.1 GI:10882238
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 949)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9172 row: m column: 12
High quality sequence stop: 656.
Location/Qualifiers
1..949
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3979187"
/clone_1ib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 296 a 200 c 251 g 202 t

ORIGIN

Query Match 36.7%; Score 66.4; DB 11; Length 949;
Best Local Similarity 64.1%; Pred. No. 3.3e-05;
Matches 100; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

DB 26 ACAGACTCTATGAAAAAGATGTGTTGTTCCATCCGCGACGAGAGAGCT 85
OY 72 ACACGGCTATATGAAACCCAAATGTATATCCAGACAAAATACAGACAGCTGGGCC 131
DB 86 TCTATATGTTTGCACACATGTGATTCATATAGTTGAATTCGTACAGACAAAAC 145
OY 133 TCCATATATGTTATGCAAAACCTCAATCTGTGTGTAATTTGCACAGAAATAACAC 191
DB 146 AACAAAGAGCTCAACAAAGCAAAATCAAAATGC 181
OY 192 AATGCAGAAAGCCCAAACTACAAAACAAATGTGC 227

RESULT 12
BG173173 972 bp mRNA EST 06-FEB-2001
LOCUS BG173173

DEFINITION 60233553F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4458709 5',
mRNA sequence.
ACCESSION BG173173
VERSION BG173173.1 GI:12679876
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 972)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10257 row: 1 column: 14
High quality sequence stop: 534.
Location/Qualifiers
1..972
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4458709"
/clone_1ib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 284 a 278 c 208 g 202 t

ORIGIN

Query Match 32.7%; Score 59.2; DB 11; Length 972;
Best Local Similarity 64.7%; Pred. No. 0.0012;
Matches 88; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

DB 126 TGTGATTCATATGTTGAATTCGTACAGAACAAACACAAAGAGCTCAACAA 165
OY 75 CCTCATTTCTGTGTGTTGATTTTGACAGAAATAACACAAATGACAGAGCCCAACT 134
DB 166 GCAAAATCAAAATGC 181
OY 135 ACNAAACAAATGTGC 150

RESULT 13
AL602490 602 bp mRNA EST 14-AUG-2001
LOCUS AL602490
DEFINITION DKFZP686P013.1 686 (synonym: hlc3) Homo sapiens cDNA clone
DKFZP686P013 5', mRNA sequence.
ACCESSION AL602490
VERSION AL602490.1 GI:1516596
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 602)
AUTHORS Ansoorge, W., Wilkner, U., Mewes, W., Well, B. and Wiemann, S.

TITLE EST (Ansoerge, W., Winkner, U., Mewes, H.W., Weill, B. and Wiemann, S.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ansoerge W

FEATURES
source
1. 602
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686P0313"
/clone_id="686 (synonym: hlcc3)"
/issue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplex2; Site_1: Sf1A; Site_2: Sf1B;
cDNA-collection"
BASE COUNT 178 a 135 c 126 g 163 t
ORIGIN

Query Match 32.5%; Score 58.8; DB 10; Length 602;
Best Local Similarity 62.0%; Pred. No. 0.0017;
Matches 93; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 32 CTCATGAAAAAGATGGATGTTCTCCATCCCGCAGAGGAGAGAGGCTTCTATA 91
DB 217 CTTTAAAGAGAGATGACCTTATCTACATACACCAAAATGCACTGCATCAATA 276
QY 92 ATGTTGCAACAACATGTTGATCTATAGTGAATCTGTACAGAACAAACACCAACA 151
DB 277 ATGTTGATGCAAAAGTTATCTCTGCTGTTGATTTGTAATGCAAGTACCAACCA 336
QY 152 GAAGCTCCAAACAAACAAATCAAAATGC 181
DB 337 GAAGCACCACCAACCTACAGACCAAGATGTC 366

RESULT 14
AL598508 758 bp mRNA EST 14-AUG-2001
LOCUS DKFZp313M1220.t1.313 (synonym: hlcc2) Homo sapiens cDNA clone
DEFINITION DKFZp313M1220 5', mRNA sequence.
ACCESSION AL598508
VERSION AL598508.1 GI:15161199
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 758)
AUTHORS Diesterhoef, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann
S.
TITLE EST (Diesterhoef, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Diesterhoef A

Am Kiofepspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No 51 sequence available.

This clone (DKFZp313M1220) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1. 758
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp313M1220"
/clone_id="313 (synonym: hlcc2)"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplex2; Site_1: Sf1A; Site_2: Sf1B;
cDNA-collection"
BASE COUNT 239 a 162 c 159 g 194 t 4 others
ORIGIN

Query Match 32.5%; Score 58.8; DB 10; Length 758;
Best Local Similarity 62.0%; Pred. No. 0.0016;
Matches 93; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 32 CTCATGAAAAAGATGGATGTTCTCCATCCCGCAGAGGAGAGAGGCTTCTATA 91
DB 337 CTTTAAAGAGAGATGACCTTATCTACATACACCAAAATGCACTGCATCAATA 430
QY 92 ATGTTGCAACAACATGTTGATCTATAGTGAATCTGTACAGAACAAACACCAACA 151
DB 431 ATGTTGATGCAAAAGTTATCTCTGCTGTTGATTTGTAATGCAAGTACCAACCA 490
QY 152 GAAGCTCCAAACAAACAAATCAAAATGC 181
DB 491 GAAGCACCACCAACCTACAGACCAAGATGTC 520

RESULT 15
BI342391 580 bp mRNA EST 30-JUL-2001
LOCUS 370140 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BI342391
ACCESSION BI342391
VERSION BI342391.1 GI:15035680
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
Fahrenkrug, S.C., Treking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.
and Keefe, J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
EST discovery (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@meat.marc.usda.gov
Single pass sequencing. Bases called and alt. trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACAT
BACKWARD: GTTTCGCCAGTCACACG
Plate: 111 row: K column: 23
Seq primer: ATTAGGTGACACTATG.
Location/Qualifiers
1. 580
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_id="MARC 2P1G"
/issue_type="pooled"
/lab_host="DH10B"

is Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2002, 20:52:55 ; Search time 521.98 Seconds
(without alignments)
78.533 Million cell updates/sec

Title: US-09-049-696-4

Perfect score: 181
Sequence: 1 CAAGAATGCATTCATCAATA.....ACAAGCAAAATCAAAATGC 181

Scoring table: IDENTITY_NUC
Gapox 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
c 1	36.8	20.3	5156	2 US-09-091-432-3	Sequence 3, Appl
c 2	35	19.3	1794	3 US-08-012-515A-13	Sequence 13, Appl
c 3	35	19.3	1794	3 US-08-360-144A-13	Sequence 13, Appl
c 4	35	19.3	1794	5 PCT-US95-06722-13	Sequence 13, Appl
c 5	32.6	18.0	1534	1 US-08-300-903A-6	Sequence 6, Appl
c 6	32.2	17.8	1117	4 US-09-247-373B-33	Sequence 33, Appl
c 7	31.2	17.2	1401	1 US-08-785-066-1	Sequence 1, Appl
c 8	31.2	17.2	1401	3 US-09-007-355-1	Sequence 1, Appl
c 9	31.2	17.2	1401	3 US-08-913-489-1	Sequence 1, Appl
c 10	31	17.1	1838	2 US-09-091-432-1	Sequence 1, Appl
c 11	30.4	16.8	3414	1 US-07-973-320-3	Sequence 1, Appl
c 12	30.2	16.7	974	2 US-08-504-459-13	Sequence 13, Appl
c 13	29.6	16.4	1743	2 US-08-591-079-1	Sequence 1, Appl
c 14	29.2	16.1	7859	1 US-07-854-596B-4	Sequence 4, Appl
c 15	29.2	16.1	7859	2 US-08-450-905B-15	Sequence 15, Appl
c 16	29.2	16.1	7859	3 US-07-982-759F-15	Sequence 15, Appl
c 17	29	16.0	621	2 US-08-951-822-6	Sequence 6, Appl
c 18	29	16.0	1454	2 US-08-713-000-7	Sequence 7, Appl
c 19	29	16.0	1454	2 US-08-975-316-7	Sequence 7, Appl
c 20	29	16.0	1454	2 US-09-211-710-7	Sequence 7, Appl
c 21	29	16.0	1474	4 US-08-975-316-71	Sequence 71, Appl
c 22	29	16.0	2280	3 US-08-813-150-1	Sequence 1, Appl
c 23	29	16.0	90050	4 US-09-245-041-5	Sequence 5, Appl
c 24	28.8	15.9	729	1 US-08-231-342-24	Sequence 24, Appl
c 25	28.8	15.9	729	1 US-08-231-342-25	Sequence 25, Appl
c 26	28.8	15.9	1366	1 US-08-231-342-22	Sequence 22, Appl
c 27	28.8	15.9	3414	1 US-07-973-320-1	Sequence 1, Appl

28.8	15.9	5892	3 US-08-755-587-27	Sequence 27, Appl
28.8	15.9	7240	3 US-08-755-587-15	Sequence 15, Appl
28.8	15.9	11283	2 US-08-603-753D-3	Sequence 3, Appl
28.8	15.9	11283	3 US-09-099-753-3	Sequence 3, Appl
28.8	15.9	11283	4 US-08-986-106-3	Sequence 3, Appl
28.8	15.9	11385	2 US-08-639-501-1	Sequence 1, Appl
28.8	15.9	11385	3 US-09-044-946-1	Sequence 1, Appl
28.8	15.9	11385	3 US-09-044-908-1	Sequence 1, Appl
28.6	15.8	467	2 US-08-841-349-18	Sequence 18, Appl
28.4	15.7	1215	4 US-09-222-851-8	Sequence 8, Appl
28.2	15.6	1134	3 US-09-248-335-29	Sequence 29, Appl
28	15.5	2209	1 US-08-514-014-1	Sequence 1, Appl
28	15.5	2209	2 US-08-833-823-1	Sequence 1, Appl
28	15.5	3214	1 US-08-484-105-17	Sequence 17, Appl
28	15.5	3214	1 US-08-484-106-17	Sequence 17, Appl
28	15.5	6143	1 US-08-612-521-3	Sequence 3, Appl
27.8	15.4	1368	3 US-08-874-563-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-091-432-3/c
Sequence 3, Application US/09091432
Patent No. 5981837

GENEAL INFORMATION:
APPLICANT: Chapple, Clint

TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
FILE REFERENCE: 7024-325

CURRENT APPLICATION NUMBER: US/09/091,432

EARLIER FILING DATE: 1998-06-18

EARLIER APPLICATION NUMBER: PCT/US96/20094

EARLIER FILING DATE: 1996-12-19

EARLIER APPLICATION NUMBER: US 60/009,119

EARLIER FILING DATE: 1995-12-22

EARLIER APPLICATION NUMBER: US 60/013,388

EARLIER FILING DATE: 1996-03-14

NUMBER OF SEQ ID NOS: 3

SOFTWARE: Microsoft Word 2.0C

SEQ ID NO 3

LENGTH: 5156

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-091-432-3

Query Match 20.3%; Score 36.8; DB 2: Length 5156;

Best Local Similarity 50.6%; Pred. No. 0.036;

Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

US-09-091-432-3
US-09-012-515A-13
Sequence 13, Application US/09012515A
Patent No. 6127521
GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
APPLICANT: Chiu, Maria Isabel
APPLICANT: Cottarel, Guillaume

APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,515A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/360,144
FILING DATE: 20-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1794 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1686
US-09-012-515A-13

Query Match 19.3%; Score 35; DB 3; Length 1794;
Best Local Similarity 54.2%; Pred. No. 0.09;
Matches 71; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 47 TGTGATTTGTTCCCAATCCCGCAGAGGCGTTCTATATAGTTTGCACACAT 106
DB 1372 TGGGGGTTTGAATTCACAAAGCGTTGGCTGATCAACGGGTATACGTTCCACAA 1431

QY 107 GTTGATTTCTATAGTTGAATTTCTGTACAGAACAAACCAACAAGAAGCTCCAAAGAG 166
DB 1432 CTCACACTGCAGAAATATTATGCGAGAGACAGATTGACGAAAAAGAAAGCTGTAAAGATTG 1491

QY 167 CAAATCAAAA 177
DB 1492 CAAAGCAAAA 1502

RESULT 3
US-08-360-144A-13
Sequence 13, Application US/08360144A
Patent No. 6150137
GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
APPLICANT: Chiu, Maria Isabel
APPLICANT: Cottarel, Guillaume
APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,144A
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1794 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1686
US-08-360-144A-13

Query Match 19.3%; Score 35; DB 3; Length 1794;
Best Local Similarity 54.2%; Pred. No. 0.09;
Matches 71; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 47 TGTGATTTGTTCCCAATCCCGCAGAGGCGTTCTATATAGTTTGCACACAT 106
DB 1372 TGGGGGTTTGAATTCACAAAGCGTTGGCTGATCAACGGGTATACGTTCCACAA 1431

QY 107 GTTGATTTCTATAGTTGAATTTCTGTACAGAACAAACCAACAAGAAGCTCCAAAGAG 166
DB 1432 CTCACACTGCAGAAATATTATGCGAGAGACAGATTGACGAAAAAGAAAGCTGTAAAGATTG 1491

QY 167 CAAATCAAAA 177
DB 1492 CAAAGCAAAA 1502

RESULT 4
PCT-US83-06722-13
Sequence 13, Application PC/TUS9506722
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Immunosuppressant Target Proteins
NUMBER OF SEQUENCES: 25
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06722
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 20-DEC-1994
INFORMATION FOR SEQ ID NO: 13:


```

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,066
FILING DATE: 17-JAN-1997
CLASSIFICATION: 432
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9601099.6
FILING DATE: 19-JAN-1996
APPLICATION NUMBER: 9622617.0
FILING DATE: 30-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gilmli, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31355-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1401 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-785-066-1

Query Match      17.2%; Score 31.2; DB 1; Length 1401;
Best Local Similarity 50.7%; Pred. No. 1.2;
Matches 75; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 13 ATTCATTAAGTAACGAGCTCTATGAAGAAAGATGAGTTCCTCCATCCGCCCA 72
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1167 AATGATTAATTTAAACAGATTATCAATTTTCAGCGAGTTTAGTGACGTTAAA 1226
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 73 GACGAGAGAGCTTCTATTAATGTTTGACACAACATGTTGATTCTATAGTTCTGTAC 132
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1227 ATCTAAATAATGACAGATGATTAATGATGAAAGATGTAATAATTAATCGAAGACGTAA 1286
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 133 AGAACAAACCAACAACAAGAGCTCCA 160
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1287 TGAAGCAAGGAAACAAAGATTGCA 1314
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-09-007-355-1
Sequence 1, Application US/09007355
Patent No. 6057138
GENERAL INFORMATION:
APPLICANT: Hodgson, John
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6057138el tRNA Synthetase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/007,355
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/785,066
FILING DATE: 17-JAN-1997
APPLICATION NUMBER: 9601099.6
FILING DATE: 19-JAN-1996
APPLICATION NUMBER: 9622617.0
FILING DATE: 30-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gilmli, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31355-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1401 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-007-355-1

Query Match      17.2%; Score 31.2; DB 3; Length 1401;
Best Local Similarity 50.7%; Pred. No. 1.2;
Matches 75; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 13 ATTCATTAAGTAACGAGCTCTATGAAGAAAGATGAGTTCCTCCATCCGCCCA 72
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1167 AATGATTAATTTAAACAGATTATCAATTTTCAGCGAGTTTAGTGACGTTAAA 1226
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 73 GACGAGAGAGCTTCTATTAATGTTTGACACAACATGTTGATTCTATAGTTCTGTAC 132
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1227 ATCTAAATAATGACAGATGATTAATGATGAAAGATGTAATAATTAATCGAAGACGTAA 1286
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 133 AGAACAAACCAACAACAAGAGCTCCA 160
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1287 TGAAGCAAGGAAACAAAGATTGCA 1314
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
US-08-913-489-1
Sequence 1, Application US/08913489
Patent No. 6080717
GENERAL INFORMATION:
APPLICANT: Hodgson, John
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6080717el tRNA Synthetase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
```


APPLICATION NUMBER: US/08/913,489
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/785,066
FILING DATE:
APPLICATION NUMBER: 9622617.0
FILING DATE: 30-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31355-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1401 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-913-489-1

Query Match 17.2%; Score 31.2; DB 3; Length 1401;
Best Local Similarity 50.7%; Pred. No. 1.2;
Matches 75; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 13 ATTCATAAAGTACAGAGCTGTATGAAAAGATGTGAGTTGTTCCAAATCCGCCA 72
DB 1167 AATTCATTAATTAAGGAGCTTATCAAAATTTTCAGCGATGTTTACGTCACCGTTAA 1226
QY 73 GACGGAGAGGCTTCTATTAAGTTGCAACATGTGTCTATGCTGATTCCTGATC 132
DB 1227 ATCTTAATTAAGTACAGATGATTTGATGAGATGTTGAATAATTAATGACGCTAA 1286
QY 133 AGAACAAAACCAACAAGAGCTCCA 160
DB 1287 TGAACCAAGAAAACAAAGATTTGCA 1314

RESULT 10
US-09-091-432-1/C
Sequence 1, Application US/09091432
Patent No. 5981837
GENERAL INFORMATION:
APPLICANT: Chapple, Clint
TITLE OF INVENTION: A Method For Regulation Of Plant Lipidn Composition
FILE REFERENCE: 7024-325
CURRENT APPLICATION NUMBER: US/09/091,432
EARLIER FILING DATE: 1998-06-18
EARLIER APPLICATION NUMBER: PCT/US96/20094
EARLIER FILING DATE: 1996-12-19
EARLIER APPLICATION NUMBER: US 60/009,119
EARLIER FILING DATE: 1995-12-22
EARLIER APPLICATION NUMBER: US 60/013,388
EARLIER FILING DATE: 1996-03-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Microsoft Word 2.0C
SEQ ID NO 1
LENGTH: 1838
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-091-432-1

Query Match 17.1%; Score 31; DB 2; Length 1838;
Best Local Similarity 56.3%; Pred. No. 1.4;
Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
QY 76 GGAGAGGCTTCTATTAAGTTGCAACATGTGATTTCTATGATTTCTGACAGA 135

DB 1819 GCACATGCTTAATTAATTAATGAAAAAATCTTATTTTACGTACATTCATCAAG 1760
QY 136 ACAAAACACAAACAAGAGCTCCAAACAAGCAAAATCAAAA 178
DB 1759 GAAATCCCATTAAGAAAAAAGGCAACAAGA 1717
RESULT 11
US-07-973-320-3
Sequence 3, Application US/07973320
Patent No. 5286486
GENERAL INFORMATION:
APPLICANT: Payne, Jewel M.
APPLICANT: Fu, Jenny M.
TITLE OF INVENTION: No. 5286486: Bacillus thuringiensis Gene
TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/973,320
FILING DATE: 19921106
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/788,638
FILING DATE: 6-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA68.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-5800
TELEFAX: 904-375-8100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3414 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: kumamotoensis
INDIVIDUAL ISOLATE: HD867
IMMEDIATE SOURCE:
LIBRARY: lamdagem (TM)-11 library of J.M. Fu
CLONE: 867
US-07-973-320-3

Query Match 16.8%; Score 30.4; DB 1; Length 3414;
Best Local Similarity 57.3%; Pred. No. 2.6;
Matches 55; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 171 CAGACGAGAGAGGCTTCTATTAAGTTGCAACAAGTGTGATTCATGATTCGT 130
DB 540 CAGATGAAATTAATCGATTTAGAACACCAACCTGTTGATTTATGATGAGCTTAT 999
QY 131 ACAGAACAAACCAACAAGAGCTCCAAACAG 166


```

; NUMBER OF SEQUENCES: 10

```

OTHER INFORMATION: /note= "Shigella flexneri 1pab"

ZIP: 60606


```

APPLICATION NUMBER: GB 9221587.0
FILING DATE: 14-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, HOLIE L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 102.378.120DV-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-526-6110
TELEFAX: 617-526-5000
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 7859 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA
US-08-450-905B-15

Query/Match          16.1%; Score 29.2; DB 2; Length 7859;
Best Local Similarity 54.7%; Pred. No. 7.8;
Matches 58; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

UY 73 GACGAGAGAGCCTCTATATGTTTGGCACACATGTTGATCTATAGTTGAATTCTGTAC 132
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5530 GCCGAGAGAGTTAGAAAATCTTGCTTTAAAAAGATTCTCTTTTATGATATTTGTAC 5609
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

UY 153 AGAACAAATCCACAAACAAGAAGCTCCAAACAGCAATCAAAA 178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5630 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5655

Search completed: April 3, 2002, 20:53:01
Job time: 41739 sec

```


This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:17:19 ; Search time 1321.64 Seconds
(without alignments)
117.412 Million cell updates/sec

Title: US-09-049-696-4

Perfect score: 181
Sequence: 1 CAAAAGATGCACATTCAATA.....ACAAGCAAAATCAAAAATGTC 181

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: N.Geneseq.1101.*

1: /SIDSL/gcgdata/geneseq/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	181	100.0	2854	22	AAH34879 Human colon cancer
2	181	100.0	2854	22	AAF81287 Human secreted pro
3	181	100.0	3111	20	AAZ09840 Human membrane spa
4	180.6	99.8	3109	22	AAH35019 Human colon cancer
5	179.4	99.1	2742	22	AAH46102 Human ICAC-1 coding
6	179.4	99.1	2745	20	AAF81927 Human ICAC-1 nuc1
7	179.4	99.1	2825	22	AAH46124 Human ICAC1 CDNA,
8	179.4	99.1	2867	22	AAH33285 Human colon cancer
9	179.4	99.1	3311	22	AAI29502 C902P determined c
10	94.8	52.4	2739	22	AAH46101 Mouse Gob-5 coding
11	94.8	52.4	2843	22	AAH46120 Mouse Gob-5 CDNA,

ALIGNMENTS

12	94.8	52.4	2931	20	AAF81925 Murine ICAC-1 nuc
13	80.8	44.6	2616	21	AA64335 Clone 251688 of a
14	80.8	44.6	3265	21	AA65095 Membrane-bound pro
15	80.8	44.6	3265	22	AA92092 Human PRO1124 CDNA
16	80.8	44.6	3265	22	AAH4241 Human PRO1124 (UNG)
17	63	34.8	307	22	AAH72076 Human cervical can
18	63	34.8	333	22	AAH71005 Human cervical can
19	63	34.8	423	22	AAH69200 Human cervical can
20	58.8	32.5	2454	22	AA682881 Human tumor-associ
21	58.8	32.5	2784	20	AA224658 Human lung tumor a
22	58.8	32.5	2784	20	AA65897 Human lung cancer-
23	58.8	32.5	3362	20	AA24657 Human lung tumor a
24	58.8	32.5	3362	21	AA65896 Human lung cancer-
25	58.8	32.5	3951	20	AA24653 Human lung tumor a
26	58.8	32.5	3951	21	AA65892 Human lung cancer-
27	58.8	32.5	8031	21	AA65950 Human lung cancer-
28	58.8	30.7	3190	20	AAH81926 Human ICAC-2 nuc1
29	58.8	21.2	5156	19	AAV10238 Arabidopsis ferula
30	38.4	21.0	3156	21	AA251625 Human membrane cha
31	36.8	20.3	5156	18	AA90505 Arabidopsis thalia
32	36.8	20.3	5156	21	AA76122 Arabidopsis ferula
33	36.2	20.0	5960	21	AA70105 Plasmidum falcipla
34	35	19.3	1794	17	AA73873 Candida RAPT1 CDNA
35	34	18.8	936	22	AA658252 Oligonucleotide D1
36	34	18.8	936	22	AA658254 Oligonucleotide D1
37	34	18.8	936	22	AA658257 Oligonucleotide D1
38	34	18.8	936	22	AA658259 Oligonucleotide D2
39	34	18.8	936	22	AA658262 Oligonucleotide D2
40	34	18.8	936	22	AA658255 Oligonucleotide D1
41	33.2	18.3	936	22	AA658252 Oligonucleotide D1
42	33.2	18.3	936	22	AA658254 Oligonucleotide D1
43	33.2	18.3	936	22	AA658257 Oligonucleotide D1
44	33.2	18.3	936	22	AA658259 Oligonucleotide D2
45	33.2	18.3	936	22	AA658262 Oligonucleotide D2

RESULTS

1	AAH34879	standard; CDNA; 2854 BP.
2	AAH34879	
3	AAH34879	
4	AAH34879	
5	AAH34879	
6	AAH34879	
7	AAH34879	
8	AAH34879	
9	AAH34879	
10	AAH34879	
11	AAH34879	
12	AAH34879	
13	AAH34879	
14	AAH34879	
15	AAH34879	
16	AAH34879	
17	AAH34879	
18	AAH34879	
19	AAH34879	
20	AAH34879	
21	AAH34879	
22	AAH34879	
23	AAH34879	
24	AAH34879	
25	AAH34879	
26	AAH34879	
27	AAH34879	
28	AAH34879	
29	AAH34879	
30	AAH34879	
31	AAH34879	
32	AAH34879	
33	AAH34879	
34	AAH34879	
35	AAH34879	
36	AAH34879	
37	AAH34879	
38	AAH34879	
39	AAH34879	
40	AAH34879	
41	AAH34879	
42	AAH34879	
43	AAH34879	
44	AAH34879	
45	AAH34879	

PS Claim 1; Page 3462-3463; 9803bp; English.

XX AAB32943 to AAB37195 and AAB73514 to AAB77788 represent human colon

CC cancer-associated nucleic acid molecules (N) and proteins (P), where

CC the proteins are collectively known as colon cancer antigens. The colon

CC cancer antigens have cytostatic activity and can be used in gene

CC therapy and vaccine production. N and P may be used in the prevention,

CC diagnosis and treatment of diseases associated with inappropriate P

CC expression. For example, N and P may be used to treat disorders

CC associated with decreased expression by rectifying mutations or deletions

CC in a patient's genome that affect the activity of P by expressing P.

CC Alternatively, N may be used to produce the colon cancer-associated Ps,

CC by inserting the nucleic acids into a host cell and culturing the cell

CC to express the proteins. N and P can be used in the prevention, diagnosis

CC and treatment of colorectal carcinomas and cancers. AAB37196 to AAB37204

CC and AAB77789 represent sequences used in the exemplification of the

CC present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were

CC missing at time of publication, meaning no sequences are present for

CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX

SO Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other.

Query Match 100.0%; Score 181; DB 22; Length 2854;

Best Local Similarity 100.0%; Pred. No. 2,9e-42;

Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAAGATGCATTCATTAAGTACAGAGCTCTATGCAAAAAGATGTGATTTGTTCT 60

DB 655 caaaagatgcattcatcaataaagtaacagagctctatgaaagagatgtgagttgtct 714

OY 61 CCAATCCCGCCAGACGAGAGGCTTCTATATGTTTCACAACTGTTGATTTATAGT 120

DB 715 ccaatcccgccagacgagagagaggtctcattatgtttcacacactgtgtatctatgt 774

OY 121 TGAATTCGTACAGACAAACACACAAAGAGCTCCAAACAGCAAAATCAAAAATG 180

DB 775 tgaattctgtacagacaacacacaaagaagctccaacaagaacaaatcaaaaatg 834

OY 181 C 181

DB 835 c 835

RESULT 2

AAF81787

ID AAF81787 standard; cDNA; 2854 BP.

XX

AC AAF81787;

XX

DT 12-JUN-2001 (first entry)

XX

DE Human secreted protein gene 1 SEQ ID NO:11.

XX

KW Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;

KW dermatological; immunosuppressive; anti-inflammatory; anti-HIV;

KW immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;

KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;

KW antialzheimer; antiparkinsonian; antimicrobial; vulnerary; gene therapy;

KW immune disorder; hyperproliferative disorder; cardiovascular disease;

KW cancer; angiogenic disorder; neurological disorder; infectious disease;

KW wound healing; regeneration; chemotaxis; chromosome 1; ss.

XX

OS Homo sapiens.

XX

PN W0200112775-A2.

XX

PD 22-FEB-2001.

XX

PF 16-AUG-2000; 2000MO-US22325.

XX

PR 17-AUG-1999; 9905-0149182.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;

PI Blise CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;

XX

DR WPI; 2001-147550/15.

DR P-PSDB; AAB74733.

XX

PT Nucleic acids encoding 25 human secreted polypeptides, useful for

PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's

PT disease and diabetic retinopathy -

XX

PS Claim 1; Page 441; 485pp; English.

XX

CC AAB81787 to AAF81817 encode the human secreted proteins given in AAB74733

CC to AAB74772. Human secreted proteins can have activities based on the

CC tissues and cells they are expressed in. Example of activities include:

CC immunomodulatory; antisclerotic; dermatological; immunosuppressive;

CC anti-inflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;

CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;

CC anticonvulsant; antialzheimer; antiparkinsonian; antimicrobial; and

CC vulnerary. Human secreted proteins can be used in gene therapy and

CC vaccine. Human secreted protein nucleotide sequences (NMI) and proteins

CC (PSP1) may be used in the prevention, diagnosis and treatment of diseases

CC associated with inappropriate polypeptide expression. For example, NMI

CC and PSP1 may be used to treat disorders associated with decreased

CC expression by rectifying mutations or deletions in a patient's genome

CC that affect the activity of proteins by expressing inactive proteins or

CC to supplement the patient's own production of polypeptides. Disorders that

CC may be prevented, diagnosed and/or treated include immune disorders,

CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,

CC angiogenic disorders, neurological disorders, infectious diseases and/or

CC for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to

CC AAB81786 and AAB74732 represent sequences used in the exemplification of

CC the present invention.

XX

SO Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other.

Query Match 100.0%; Score 181; DB 22; Length 2854;

Best Local Similarity 100.0%; Pred. No. 2,9e-42;

Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAAGATGCATTCATTAAGTACAGAGCTCTATGCAAAAAGATGTGATTTGTTCT 60

DB 655 caaaagatgcattcatcaataaagtaacagagctctatgaaagagatgtgagttgtct 714

OY 61 CCAATCCCGCCAGACGAGAGGCTTCTATATGTTTCACAACTGTTGATTTATAGT 120

DB 715 ccaatcccgccagacgagagagaggtctcattatgtttcacacactgtgtatctatgt 774

OY 121 TGAATTCGTACAGACAAACACACAAAGAGCTCCAAACAGCAAAATCAAAAATG 180

DB 775 tgaattctgtacagacaacacacaaagaagctccaacaagaacaaatcaaaaatg 834

OY 181 C 181

DB 835 c 835

RESULT 3

AZ09840

ID AZ09840 standard; cDNA; 3111 BP.

XX

AC AZ09840;

XX

DT 26-NOV-1999 (first entry)

XX

DE Human membrane spanning protein MSP-5 cDNA fragment 2.

XX

KW Membrane spanning protein; MSP; human; treatment; diagnosis; prevention;

Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
expression inhibition; antisense therapy; gene therapy;
chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
ds.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..2742

XX FT /*tag= a

XX FT /product= "Human CLCA1"

XX FT /note= "No stop codon given in the specification"

XX PN WO200138530-A1.

XX PD 31-MAY-2001.

XX PF 22-NOV-2000; 2000WO-JP08232.

XX PR 24-NOV-1999; 99JP-0333479.

XX PR 27-APR-2000; 2000JP-0127589.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Nakanishi A, Morita S;

XX DR WPI: 2001-355935/37.

XX DR P-PSDB; AAB73716.

XX PT New antisense nucleotide, useful for treatment and prevention of

XX PS bronchial asthma and chronic obstructive pulmonary disease

XX PS Claim 3; Page 82-84; 104pp; Japanese.

CC The invention relates to an antisense nucleotide targeted to the mouse

CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,

CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also

CC relates to an antibody specific for the Gob-5 protein, medical and

CC diagnostic compositions containing the antisense nucleotide or the

CC antibody, and methods and kits for screening for compounds which inhibit

CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.

CC The antisense oligonucleotides and antibody are therefore useful for the

CC treatment and prevention of bronchial asthma and chronic obstructive

CC pulmonary disease. The present sequence represents the human CLCA1

CC gene coding sequence.

XX SQ Sequence 2742 BP; 833 A; 616 C; 623 G; 670 T; 0 other;

Query Match 99.1%; Score 179.4; DB 22; Length 2742;

Best Local Similarity 99.4%; Pred. No. 8.1e-42;

Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

YY 1 CAAAGATGCGCATTCATTAAGTAACGAGCTCTATGAAAAGAGTGTGTTGTTCT 60

Db 621 caaagatgcatcattcaataaagtaacgagctctatgaaaagagatgtgtgtct 680

YY 61 CCAATCCCGCCAGACGAGAGGCTTCTATATGTTTGACACAGATGTTGATTATAGT 120

Db 681 ccaatcccgccagacgagagcttctataatgtttgacacaaatgtgtctatagt 740

YY 121 TGAATTCTGTACAGACAAACCAACAAAGAGCTCCAAACAGCAAAATCAAAAATG 180

Db 741 tgaattctgtacagacaaacacacaaagaagctccaaacaaatcaaaaatg 800

YY 181 C 181

Db 801 c 801

RESULT 6

AAF81927

AAE81927 standard; cDNA; 2745 BP.

AC AAF81927;

XX 13-JUN-2001 (first entry)

XX DE Human ICACC-1 nucleotide sequence.

XX KW ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;

XX KW Interleukin 9-induced calcium activated chloride channel; IL-9;

XX KW calcium activated chloride channel; anti-allergic; anti-asthmatic;

XX KW anti-inflammatory; immunomodulatory; cystic fibrosis;

XX KW inflammatory bowel disease; autoimmune disease; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..2745

XX FT /*tag= a

XX FT /product= "ICACC-1"

XX FT /note= "IL-9 induced calcium activated chloride channel"

XX PN WO944620-A1.

XX PD 10-SEP-1999.

XX PF 03-MAR-1999; 99WO-US04703.

XX PR 03-MAR-1998; 98US-0076815.

XX PA (NAGA-) MAGAININ PHARM INC.

XX PI Holroyd RJ, Levitt RC, Maloy WL, Louahed J, McLane M;

XX PI Nicolalde NC, Zhou Y, Dong Q;

XX DR WPI: 1999-550979/46.

XX DR P-PSDB; AAB74824.

XX PT New nucleic acid encoding calcium activated chloride channel, used to

XX PS identify, e.g. specific modulators for treating atopic allergy

XX PS Claim 1; Fig 4B; 75pp; English.

CC The present sequence encodes the human interleukin 9 (IL-9) induced

CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins

CC have anti-allergic, anti-asthmatic, anti-inflammatory and

CC immunomodulatory activities. Compounds (A) that downregulate ICACC are

CC used to alleviate asthma (or more generally atopic allergy), while those

CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and

CC inflammatory bowel disease (IBD) (or other autoimmune diseases).

CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma

CC (increased levels) or IBD (reduced levels), also for monitoring

CC treatment of these conditions. The ICACC proteins can be used:

CC (i) to raise specific antibodies (Ab), useful: (a) as immunoassay

CC reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as

CC (A) to disrupt bonding between ICACC and its ligands; and (iii) to

CC identify modulators and binding partners. ICACC polynucleotides can be

CC used to generate transgenic animals or recombinant cells, used to screen

CC for antagonists, also as a source of therapeutic antisense agents or

CC diagnostic probes (for quantifying mRNA expression, e.g. for

CC identification of modulators).

XX SQ Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;

Query Match 99.1%; Score 179.4; DB 20; Length 2745;

Best Local Similarity 99.4%; Pred. No. 8.1e-42;

Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

YY 1 CAAAGATGCGCATTCATTAAGTAACGAGCTCTATGAAAAGAGTGTGTTGTTCT 60

Db 621 caaagatgcatcattcaataaagtaacgagctctatgaaaagagatgtgtgtct 680

QY 61 CCAATCCCGCAGAGGAGGAGGCTTCTATTAATGTTTGCACAACATGTTGATTTCTACT 120
DB 681 ccaatcccgccagagagagagcttctataatgttcacacaacatgttctatagt 740
QY 121 TGAATTCTGTACAGAACAAACCAAGAGCTCCAAACCAACCAAAATCAAAAATG 180
DB 741 Tgaattctgtacagacaacaccacaagaagctccaacaagaacaaatcaaaaatg 800
QY 181 C 181
DB 801 c 801

RESULT 7
AAH46124
ID AAH46124 standard; cDNA; 2825 BP.
XX
AC AAH46124;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human CLCA1 cDNA, SEQ ID NO:26.
XX
KM Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;
KM expression inhibition; antisense therapy; gene therapy;
KM chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
SS.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 23..2767
FT CDS /tag= a
FT /product= "Human CLCA1"
FT /transl_except= (pos:476..478, aa:lys)
XX
PN WO200138530-A1.
XX
PD 31-MAY-2001.
XX
PE 22-NOV-2000; 2000MO-JP08232.
XX
PR 24-NOV-1999; 99JP-0333479.
PR 27-APR-2000; 2000JP-0127589.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Nakanishi A, Morita S;
XX
DR WPI; 2001-355935/37.
DR P-PSDB; AAB73716.
XX
PT New antisense nucleotide, useful for treatment and prevention of
PT bronchial asthma and chronic obstructive pulmonary disease -
XX
XX Example 5; Page 92-94; 104pp; Japanese.
XX
XX The invention relates to an antisense nucleotide targeted to the mouse
CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC relates to an antibody specific for the Gob-5 protein, medical and
CC diagnostic compositions containing the antisense nucleotide or the
CC antibody, and methods and kits for screening for compounds which inhibit
CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC The antisense oligonucleotides and antibody are therefore useful for the
CC treatment and prevention of bronchial asthma and chronic obstructive
CC pulmonary disease. The present sequence represents human CLCA1 cDNA.
XX
SQ Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;

Query Match 99.1%; Score 179.4; DB 22; Length 2825;
Best Local Similarity 99.4%; Pred. No. 8.2e-42;

Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CAAAGATGCACATTCATTAATTAAGTAACAGAGCTCTATGAAAAGAGTGTGTTGTC 60
DB 1643 caaagatgtcaccttcaataaagttacagacctatgaaaaagatgtgagtttttc 702
QY 61 CCAATCCCGCAGAGGAGGAGGCTTCTATTAATGTTTGCACAACATGTTGATTTCTACT 120
DB 703 ccaatcccgccagagagagcttctataatgttcacacaacatgttctatagt 762
QY 741 TGAATTCTGTACAGAACAAACCAAGAGCTCCAAACCAACCAAAATCAAAAATG 180
DB 743 Tgaattctgtacagacaacaccacaagaagctccaacaagaacaaatcaaaaatg 822
QY 181 C 181
DB 823 c 823

RESULT 8
AAH33285
ID AAH33285 standard; cDNA; 2867 BP.
XX
AC AAH33285;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:341.
XX
KM Human; colon cancer; colon cancer antigen; diagnosis; detection;
KM colorectal carcinoma; chromosome 1; ss.
XX
XX Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PE 28-SEP-2000; 2000MO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruden SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR P-PSDB; AAG73854.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1; Page 2452-2453; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated P,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
XX
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for

CC The present invention describes: colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytosolic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be

XX
XX

XX	31-MAY-2001.
PD	
XX	
PF	22-NOV-2000; 2000MO-JP08232.
XX	
XX	24-NOV-1999; 99UP-0333479.
3R	

27-APR-2000; 2000JP-0127589.
 (TAKE) TAKEDA CHEM IND LTD.
 Nakanishi A, Morita S;
 WPI; 2001-355935/37.
 P-PSDB; AAB73715.
 New antisense nucleotide, useful for treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease -
 Claim 3; Page 80-82; 104pp; Japanese.
 The invention relates to an antisense nucleotide targeted to the mouse Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart, the CLCA1 gene (coding sequence shown in AAH46102). The invention also relates to an antibody specific for the Gob-5 protein, medical and diagnostic compositions containing the antisense nucleotide or the antibody, and methods and kits for screening for compounds which inhibit the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells. The antisense oligonucleotides and antibody are therefore useful for the treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease. The present sequence represents the mouse Gob-5 gene coding sequence.
 Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other;

Query Match 52.4%; Score 94.8; DB 22; Length 2739;
 Best Local Similarity 70.8%; Pred. No. 9.4e-18;
 Matches 126; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
 4 AAGATGCATCTCAATTAAGTAACAGAGCTCTATGAAAAGATGAGTTGTTCCCA 63
 627 aaagtggttaactgacagaggaagctgacataagaacaattggtatgtaccaga 666
 64 ATCCCGCAGACGAGAGAGCTTCTATTAATGTTGGACAACATGTTGATTTAGTTGA 123
 687 tccacacacaaagagaagagcttccatcatgttaccacaaatataatcttggttga 746
 124 ATTCTGTACAGAACAAACACACAAAGAGTCCAAACAGCAAAATCAAAAATGC 181
 747 attctgtacagaaaaaatcacatcaagaagcccaaatgacaaacacacagatgc 804

RESULT 11
 AAH46120
 ID AAH46120 standard; cDNA; 2843 BP.
 AC AAH46120;
 DT 11-SEP-2001 (first entry)
 DE Mouse Gob-5 cDNA, SEQ ID NO:22.
 XX
 XX Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening; expression inhibition; antisense therapy; gene therapy; bronchial asthma; chronic obstructive pulmonary disease; antisthmatic; ss.
 OS Mus sp.
 FH Key Location/Qualifiers
 FT CDS 15..2756
 FT /*tag= a
 FT /product= "Mouse Gob-5"
 PN WO200138530-A1.
 XX 31-MAY-2001.
 XX 22-NOV-2000; 2000OMO-JP08232.
 XX

24-NOV-1999; 99JP-0333479.
 27-APR-2000; 2000JP-0127589.
 (TAKE) TAKEDA CHEM IND LTD.
 Nakanishi A, Morita S;
 WPI; 2001-355935/37.
 P-PSDB; AAB73715.
 New antisense nucleotide, useful for treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease -
 Example 1; Page 89-91; 104pp; Japanese.
 The invention relates to an antisense nucleotide targeted to the mouse Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart, the CLCA1 gene (coding sequence shown in AAH46102). The invention also relates to an antibody specific for the Gob-5 protein, medical and diagnostic compositions containing the antisense nucleotide or the antibody, and methods and kits for screening for compounds which inhibit the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells. The antisense oligonucleotides and antibody are therefore useful for the treatment and prevention of bronchial asthma and chronic obstructive pulmonary disease. The present sequence represents mouse Gob-5 cDNA.
 Sequence 2843 BP; 815 A; 705 C; 680 G; 643 T; 0 other;

Query Match 52.4%; Score 94.8; DB 22; Length 2843;
 Best Local Similarity 70.8%; Pred. No. 9.5e-18;
 Matches 126; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
 4 AAGATGCATCTCAATTAAGTAACAGAGCTCTATGAAAAGATGAGTTGTTCCCA 63
 641 aaagtggttaactgacagaggaagctgacataagaacaattggtatgtaccaga 700
 64 ATCCCGCAGACGAGAGAGCTTCTATTAATGTTGGACAACATGTTGATTTAGTTGA 123
 701 tccacacacaaagagaagagcttccatcatgttaccacaaatataatcttggttga 760
 124 ATTCTGTACAGAACAAACACACAAAGAGTCCAAACAGCAAAATCAAAAATGC 181
 761 attctgtacagaaaaaatcacatcaagaagcccaaatgacaaacacacagatgc 818

RESULT 12
 AAF81925
 ID AAF81925 standard; cDNA; 2931 BP.
 AC AAF81925;
 DT 13-JUN-2001 (first entry)
 DE Murine ICACC-1 nucleotide sequence.
 XX
 XX ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder; interleukin 9 induced calcium activated chloride channel; IL-9; calcium activated chloride channel; anti-allergic; anti-asthmatic; anti-inflammatory; immunomodulatory; cystic fibrosis; inflammatory bowel disease; autoimmune disease; ss.
 OS Mus sp.
 FH Key Location/Qualifiers
 FT CDS 8..2749
 FT /*tag= a
 FT /product= "ICACC-1"
 FT /note= "IL-9 induced calcium activated chloride channel."
 PN WO9944620-A1.
 XX 10-SEP-1999.
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 02:56:57 ; Search time 22700.8 Seconds

(Without alignments)
131.537 Million cell updates/sec

Title: US-09-049-696-4

Perfect score: 181
Sequence: 1 CAAAGATGACATTCATAA.....ACAAGCAAAATCAAAAATGC 181

Scoring table: IDENTITY_NTC
Gapop 10.0, Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_ov: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vi: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rod: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	179.4	99.1	2826	9	AF127036 Homo sapi
2	179.4	99.1	3311	6	AX193489 Sequence
3	179.4	99.1	3311	9	AF039400 Homo sapi
4	117	64.6	113764	9	HSJ651E10 Human DNA
5	115.4	63.8	35278	9	AF039401 Homo sapi
6	115.4	63.8	164891	2	AL358950 Homo sapi
7	105.8	58.5	3079	4	AF095584 Sus scrofa
8	94.8	52.4	2937	10	AB017156 Mus muscu
9	80.8	44.6	3204	9	AF127035 Homo sapi
10	80.8	44.6	3221	9	AK000072 Homo sapi
11	80.8	44.6	3265	6	AX092338 Sequence
12	74.2	41.0	1265	4	AF001264 Bos tauru
13	74.2	41.0	2820	4	AF001263 Bos tauru
14	74.2	41.0	3288	4	AF001262 Bos tauru
15	74.2	41.0	3317	4	AF001261 Bos tauru
16	72.8	40.2	3415	9	AF043976 Homo sapi
17	71.2	39.3	3022	10	AF047838 Mus muscu
18	71.2	39.3	3471	10	AF052746 Mus muscu
19	67.8	37.5	2984	4	BTU36445 Bos tauru
20	66.4	36.7	2765	10	AF115852 Mus muscu
21	66.4	36.7	3058	10	AF108501 Mus muscu
22	66.4	36.7	3137	10	BC008147 Mus muscu
23	63	34.8	307	6	AX187657 Sequence
24	63	34.8	333	6	AX186584 Sequence
25	63	34.8	423	6	AX184779 Sequence
26	58.8	32.5	2832	9	AF127980 Homo sapi
27	58.8	32.5	2970	9	AF043977 Homo sapi
28	58.8	32.5	3604	9	AB026833 Homo sapi
29	58.8	32.5	4077	6	AX054697 Sequence
30	45	24.9	140718	2	AL356270 Homo sapi
31	41.8	23.1	140718	2	AL356270 Homo sapi
32	41.8	23.1	175591	2	AC068071 Homo sapi
33	38.4	21.2	1010	5	GGCOLA2C Arabidops
34	38.4	21.2	1248	5	CHRCOLA29 Arabidops
35	38.4	21.2	1683	5	GGCO10 Arabidops
36	38.4	21.2	1691	3	CHRC1PA2M Arabidops
37	37	20.4	1011	3	DME300073 Arabidops
38	37	20.4	251199	9	HUAE000659 Homo sapi
39	36.8	20.3	5156	6	AR085703 Sequence
40	36.8	20.3	8700	8	AF068574 Arabidops
41	36.8	20.3	94695	8	AF23E13 Arabidops
42	36.8	20.3	198750	8	ATC81985 Arabidops
43	36.2	20.0	2392	9	BC007680 Homo sapi
44	36.2	20.0	9432	3	AE001375 Homo sapi
45	35.8	19.8	157165	4	AC087843 Bos tauru

ALIGNMENTS

RESULT	1	AF127036	2826 bp	mrna	PRI	10-AUG-1999
LOCUS	AF127036					
DEFINITION		Homo sapiens calcium-activated chloride channel protein 1 (CACCL1)				
KEYWORDS		mrna, complete cds.				
ACCESSION		AF127036				
VERSION		AF127036.1				
KEYWORDS		GI:4585468				
SOURCE						
ORGANISM		human				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS		Aguel, M., Vermet, T. and Culouscou, J. M.				
TITLE		Identification of three novel members of the calcium-dependent chloride channel (CACCL) family predominantly expressed in the digestive tract and trachea				
JOURNAL		FEBS Lett. 455 (3), 295-301 (1999)				
MEDLINE		99364503				
PUBMED		10437792				

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
2	(bases 1 to 2826)	Aguel, M. and Culooson, J.-M.	Submitted (03-FEB-1999)	Grouppe Genomique, SYNTHELABO, 10 Rue des Carrières, Neuilly-Malmaison 92500, France
source	gene	CDs	location/Qualifiers	
1. .2826	/organism="Homo sapiens"	/db_xref="taxon:9606"	/tissue_type="small intestine; colon"	
1. .2826	/gene="Caccl1"			
5. .2749	/gene="Caccl1"			
/note="povine epithelial chloride channel homolog"				
/codon_start=1				
/product="calcium-activated chloride channel protein 1"				
/protein_id="AAD25487.1"				
/db_xref="gi:4585469"				
/translation="MGPEKSSVFLLILHLLEGALSNLIOLNNNGYEIVVADPNVEDELILIOIKDMVTOASLYLEATGKRRYFNVAHLIPETMKADYADYRPLLEFYKNA				
DVIAVSPSTPNDEPRDPMQNGCEGKRRYFNLIPEDPDLAKYKAGYRQGRAFYEMAH				
LRMVDFEPTNNDKPEFYLSNGRIQAVRCAGITGTNNVAKCGGSCYTRKCFNVTGI				
YERKEEFLVQSRQTEKASIMFAQHVDSVECTQDNHKKKAPNQNCKNLRSWEVY				
LDSDSEFKTTPMTQPPNPFSLQIQIRIVCLVDLNGSGMATGNRLNLMQAOFLF				
LOVLEJSDGWMATWFDSSAAVHOSLQIQINSGSDIDTLAKRLPAASGGTSICSLRS				
PTVTRKKYPTDGEIVLITLDEEDNTISGCFNEVOSQIILHTVALGPSAAGLEIELS				
MTGLQIYASDOVONNGILDAFGALSSNGAVSRSIOLIESKGLTIONSOMNCTIVY				
DSYVKGDTLLITWTQPPRQTLIMDDPSGKCGFVVDKNTKMAVLIQIPGIAKVTWKY				
SLQASSTQTLITLTVTSRSNATLPPITYTSKTKNKTSPFPLVYVYANIRQASPTLR				
SVTALIESVNGKIVTLELDNGAGADATKDDVYSRFTTYDNGRYSVKYRALGVNA				
AAARRVPIPDLEPQGITDILKAEIHGSLINLTWTAPGDDYDHGAHKYITRISILDE				
DVRRAPVLPDLEPQGITDILKAEIHGSLINLTWTAPGDDYDHGAHKYITRISILDE				
LDRKFENESLIQVNTALIPKEANSESEVLEFKENITPENGTDLFLAIOAVDVKDLSRE				
SNIRVSLFLPDPPTPEPSPDENSAPCPNHNISTIPGHIILTKIMKMWIGELDLSIA				
BASE COUNT	875 a	623 c	632 g	696 t
ORIGIN				
Query Match	99.1%;	Score 179.4;	DB 9;	Length 2826;
Best Local Similarity	99.4%;	Pred. No. 3.7e-36;		
Matches 180;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps 0;				
1 CAAAGATGCACATTCATTAAGATACAGAGACTATGAGAAAAAGATGAGTTTGTCT				60
625 CAAAGATGCACATTCATTAAGATACAGAGACTATGAGAAAAAGATGAGTTTGTCT				684
61 CCAATCCCGCCGAGAGGAGGCTTCATATATGTTTGACAAACATGTGATTCTATGT				120
665 CCAATCCCGCCGAGAGGAGGCTTCATATATGTTTGACAAACATGTGATTCTATGT				744
121 TGAATTCGTACAGAACAAACACACACAGAACAGCTCCAAAACAAGCAAAATCAAAAATG				180
745 TGAATTCGTACAGAACAAACACACACAGAACAGCTCCAAAACAAGCAAAATCAAAAATG				804
181 c				
805 c				
RESULT 2				
AX193489	3311 bp	DNA	PAT	15-AUG-2001
DEFINITION	Sequence 1056 from Patent WO0149716.			
ACCESSION	AX193489			
VERSION	AX193489.1	GI:15211440		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
1 (bases 1 to 3311)	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			

AUTHORS	Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J., Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE	Compounds for immunotherapy and diagnosis of colon cancer and methods for their use
JOURNAL	Patent: WO 0149716-A 1056 12-JUL-2001;
FEATURES	CORIXA CORPORATION (US)
SOURCE	1. .3311 Location/Qualifiers
BASE COUNT	1028 a 692 c 742 g 849 t
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606"
Query Match	99.1%; Score 179.4; DB 6; Length 3311;
Best Local Similarity	99.4%; Pred. No. 3.7e-36;
Matches 180; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1 CAAAGATGCACATTCATTAAGTTAAACAGAGCTCATGAAAAAGGATGTGATTTGTTCT 60
DB	972 CAAAGATGCACATTCATTAAGTTAAAGTTACAGAGCTCATGAAAAAGGATGTGATTTGTTCT 1031
OY	61 CCAATCCGCCACAGACGAGGAGGCTTCTATTAATGTTTGCACAACTGTTGATTCATAGT 120
DB	1032 CCAATCCGCCACAGACGAGGAGGCTTCTATTAATGTTTGCACAACTGTTGATTCATAGT 1091
OY	121 TGAATTCGTGCAGAACAAACCCACAAAGAGCTCCAAACAGCAAAATCAAAATG 180
DB	1092 TGAATTCGTGCAGAACAAACCCACAAAGAGCTCCAAACAGCAAAATCAAAATG 1151
OY	181 C 181
DB	1152 C 1152
RESULT 3	
AF039400	
LOCUS	AF039400 3311 bp mRNA. PRI 14-DEC-1998
DEFINITION	Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA, complete cds.
ACCESSION	AF039400
VERSION	AF039400.1 GI:4009457
KEYWORDS	human.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 3311)
AUTHORS	Gruber,A.D., Eldle,R.C., Ji,H.L., Schreier,K.D., Fuller,C.M. and Pauli,B.U.
TITLE	Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of Ca2+-activated Cl- channel proteins
JOURNAL	Genomics 54 (2), 200-214 (1998)
MEDLINE	99047526
REFERENCE	2 (bases 1 to 3311)
AUTHORS	Gruber,A.D., Eldle,R. and Pauli,B.U.
TITLE	Direct Submission
JOURNAL	Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
FEATURES	Location/Qualifiers
SOURCE	1. .3311 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="1" /map="1p22-p31" /tissue-type="small intestine" 1. .3311 /gene="hCLCA1" 352. .3096 /gene="hCLCA1" /note="transmembrane glycoprotein" /codon_start=1
gene	
CDS	

/product="calcium-dependent chloride channel-1"
 /protein_id="AAC95428.1"
 /db_xref="GI:4009458"
 /translation="MGPKSSVFILILHLEGAISNSLIOLNNNGYEGIVAVADPNP
 EDFTLLOIKMVAQSLYLFEATGKREYFNVALILPEPTGKTRKAPKLEFYKNA
 DVLVAESTPPGNDPEPTQMGNGCKGRGRIHLPFTIGAKLAETGPGKAFVHMIAH
 LRMGVFEDYNDKERYLSNGIAVRCSAGITGVNKKCGGSCYTLKCTENKVTGL
 YKGEFVLQSRQTEKASIMFAQHDVSIVERTEQNHKEAPNQNCKNLSTEWEL
 RDSDFPKTPTPTQPPNPPTSLQIGRIVCILVJDKSGMATGRLNLOACGLFL
 LOYVLSGVMGVFPDASAHOSELIQINSGRDRLAKRLPAAASGGSTCSGFRSA
 FTVLRKKYPPGSEYVILLTGDENITISGPNVKGSAIITHVALGSPSAOELERLSK
 MFGCLQYASDVQVONGICIDAFGLSSNGAVSQSTOLESGTLQNSOMNGTVIV
 DSTVGKDLFLITWTPQQLIMDPSSQKGGEVVDKNTKVAIQLPGLAVNGTVY
 SLQSSQTLTLTVTSRASNALPPIITVSKTKNDKSPSPVYVANIROGASPIIRA
 SVALIESVNGKTVLLELDNGAGADARQGVSRFTYDNGRYSKVVALGCVN
 AARRVIRPOOSGALYIPGMIEDEIOMNPPEIKKDVOHOKVCEFRSSGSPVAS
 DVNAPRIPDLPEPGIOTDLKAIHGGSLINTLWAPAGDGYOHAHKYIIRTSYILD
 LROKFNESLQVNTLRALIRKEANSEVEFLFKPENITFENGTDLFIAIOVDKYLSEI
 SNIAVSLFIPQIPETPSPDTSAPCNHINHTPGIHLKIMMKWIGELQLSIA

BASE COUNT 1028 a 692 c 742 g 849 t
 ORIGIN
 Query Match 99.1%; Score 179.4; DB 9; Length 3311;
 Best Local Similarity 99.4%; Pred. No. 3.7e-36;
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CAAAGATGCAATTCATAAAGTAACAGACTCTATGAAAAAGATGAGTTTCT 60
 Db 972 CAAAGATGCAATTCATAAAGTAACAGACTCTATGAAAAAGATGAGTTTCT 1031
 OY 61 CCAATCCCGCCAGAGAGAGAGAGCTTCTATATGTTTGCACAACATGATGATCTACT 120
 Db 1032 CCAATCCCGCCAGAGAGAGAGAGCTTCTATATGTTTGCACAACATGATGATCTACT 1091
 OY 121 TGAATTCGTGTACAGAACAAACACAAAGAGCTCCAAACCAACCAAAATCAAAAAATG 180
 Db 1092 TGAATTCGTGTACAGAACAAACACAAAGAGCTCCAAACCAACCAAAATCAAAAAATG 1151
 OY 181 c 181
 Db 1152 c 1152

RESULT 4
 LOCUS HSJ651E10 113764 bp DNA PRI 27-MAY-2000
 DEFINITION Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1,
 complete sequence.
 ACCESSION AL122002
 VERSION AL122002.16 GI:8247274
 KEYWORDS HMG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 113764)
 JOURNAL Direct Submission
 Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jun 4, 2000 this sequence version replaced gi:8247030.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence has been finished according to sequence map criteria
 as follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 EMBL: SW, SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/projects/C-elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr1
 RP4-651E10 is from the library RPci-4 constructed at the Roswell
 Park Cancer Institute by the group of Pieter de Jong. For further
 details see http://nacpac.med.buffalo.edu/
 VECTOR: pcypAC2
 This sequence is the entire insert of clone RP4-651E10.
 FEATURES
 source
 1. .113764
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="p22.3-31.1"
 /clone="RP4-651E10"
 /clone_1fb="RPci-4"

BASE COUNT 34890 a 21989 c 22351 g 34534 t
 ORIGIN

Query Match 64.6%; Score 117; DB 9; Length 113764;
 Best Local Similarity 92.5%; Pred. No. 2.9e-20;
 Matches 123; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 OY 1 CAAAGATGCAATTCATAAAGTAACAGACTCTATGAAAAAGATGAGTTTCT 60
 Db 15238 CAAAGATGCAATTCATAAAGTAACAGACTCTATGAAAAAGATGAGTTTCT 15297
 OY 61 CCAATCCCGCCAGAGAGAGAGCTTCTATATGTTTGCACAACATGATGATCTACT 120
 Db 15238 CCAATCCCGCCAGAGAGAGAGCTTCTATATGTTTGCACAACATGATGATCTACT 15357
 OY 121 TGAATTCGTGTACAGAACAAACACAAAGAGCTCCAAACCAACCAAAATCAAAAAATG 183
 Db 15358 TACCTTGTCTCA 15370

RESULT 5
 LOCUS AF039401 35278 bp DNA PRI 14-DEC-1998
 DEFINITION Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) gene,
 complete cds.
 ACCESSION AF039401
 VERSION AF039401.1 GI:4009459
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 35278)
 JOURNAL Gruber A.D., Elble R.C., Ji H.L., Schreuer K.D., Fuller C.M. and
 Pauli B.U.
 Genomic cloning, molecular characterization, and functional
 analysis of human CLCA1, the first human member of the family of
 Ca2+-activated Cl- channel proteins
 Genomics 54 (2), 200-214 (1998)
 MEDLINE 99047526
 JOURNAL 2 (bases 1 to 35278)
 REFERENCE Gruber A.D., Elble R., and Pauli B.U.
 Direct Submission
 Submitted (19-DEC-1997) Department of Pathology, College of
 Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA
 JOURNAL
 TITLE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 JOURNAL
 FEATURES
 source
 1. .35278
 Location/Qualifiers


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22-p31"
1.1617
/promoter
/gene="hCLCA1"
1..33522
/gene="hCLCA1"
repeat_region
767..1265
/rpt_family="L1 transposable element"
/rpt_type="dispersed"
1587..1596
/gene="hCLCA1"
/feature="putative"
1618..1708
/gene="hCLCA1"
/number=1
join(1618..1708,1962..2383,6667..6807,6979..7126,
9696..9801,15454..15631,18591..18809,19783..20010,
22253..22427,24522..24628,26644..26859,27447..27708,
28766..28936,31840..32079,32919..33522)
/gene="hCLCA1"
/product="calcium-dependent chloride channel-1"
1962..2383
/gene="hCLCA1"
/number=2
join(2222..2383,6667..6807,6979..7126,9696..9801,
15454..15631,18591..18809,19783..20010,22253..22427,
24522..24628,26644..26859,27447..27708,28766..28936,
31840..32079,32919..33522)
/gene="hCLCA1"
/codon_start=1
/product="calcium-dependent chloride channel-1"
/protein_id="AAC95429.1"
/db_xref="GI:4009460"
/translacion="MGPKSSVFLLIHLEGALSNSLQLNNGYRGIVADPNP
EDFTLQIKDMYTOASLYFEATGRKFRYKNAVILIPETKTKADYRKLRTKNA
DLVAESTPGEDEPYEQMNGCEGKERHLPDFIAGKKLAEYGPQAKAFYEMH
LRWGVDEYNNDEKFLSNRIQAVRCSAIGTNVVKCOGSCYTRKCTFNVTGL
YKGCFFVLQSRQTEKASIMPAQHVDISVECTEQNHKKAAPKONOKMLRSTWEV
RSEDEFKTTPMTQPPNPTEFSLQICORIVCLVDKSGSAMGNRLNRLNQGOLF
LQTELSWAGYTFDSAAVQSELQINSQSDPDLAKRLPAASGSGTSCISLBSA
FYIRKYPDQSEIVLLTDEGNTISGCTNEKYSQSAIIHYVALGSAQOEELSK
MTGGLQYADQVONNELIDAFGLSNGKAVSQRSLQSKGLTLONSQMMNTVIV
DSTVGKDTFLITWTQPPQILMLDPGQOGGVQVQDKRMVLIQIPGAKGTMWY
SLQASQDTLTLVTSRASNAITLPITVTSKNTKSFPSPLVYVARIQASGILRA
SVALLSVNGKVTLELLDNGAGADTKDGYRSRFTYDNGRYSVRVAGVNA
AARRVYIPQOSGALYIPGWIENDEIONPRLPNKDDVOHKQVCESTSGSFVNS
DVNAPIPDLFPPGQITDLNAETHGSLIMLTARBDQDHRGNAHYIIRISIID
LRKFNESLOVNTTALIPKANSEEVFLPENTTFENGIDFLTAIDAVKVDLKSI
SNIAVSLFIPQTPPTSPDSTAPCPNHIINSTIPGIHLIKIMKWIQELQLSIA
"
exon
6667..6807
/gene="hCLCA1"
/number=3
6979..7126
/gene="hCLCA1"
/number=4
9696..9801
/gene="hCLCA1"
/number=5
15454..15631
/gene="hCLCA1"
/number=6
18591..18809
/gene="hCLCA1"
/number=7
19783..20010
/gene="hCLCA1"
/number=8
22253..22427
/gene="hCLCA1"
/number=9

```

```

exon
24522..24628
/gene="hCLCA1"
/number=10
26644..26859
/gene="hCLCA1"
/number=11
27447..27708
/gene="hCLCA1"
/number=12
28766..28936
/gene="hCLCA1"
/number=13
31840..32079
/gene="hCLCA1"
/number=14
32919..33522
/gene="hCLCA1"
/number=15
exon
BASE COUNT 11245 a 6870 c 7230 g 9933 t
ORIGIN

```

```

Query Match 63.8% Score 115.4; DB 9; Length 35278;
Best Local Similarity 91.7% Pred. No. 8e-20;
Matches 122; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 1 CAAAAGATGCATTCATTAAGTAACAGACGCTGTATGAGAAAGATGCTGTTGTTCT 60
|||||
Db 15517 CAAAAGATGCATTCATTAAGTAACAGACGCTGTATGAGAAAGATGCTGTTGTTCT 15576
|||||
QY 61 CCAATCCCGCAGACGAGAGAGGCTTCTATATATGTTTGACACACATGTTGATTTAGT 120
|||||
Db 15577 CCAATCCCGCAGACGAGAGAGGCTTCTATATATGTTTGACACACATGTTGATTTAGT 15636
|||||
QY 121 TGAATTCGTCTCA 133
Db 15637 TACCTGTCTCA 15649

```

```

RESULT 6
AL358950/c
LOCUS
DEFINITION
AL358950
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 164891)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
Project Information
Center project name: BA44AC12
Summary Statistics
Sequencing program: XGAP4; Version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least 040
Consensus quality: 162610 bases at least 030
Insert size: 163391; sum-of-ctrls

```

```

Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
Project Information
Center project name: BA44AC12
Summary Statistics
Sequencing program: XGAP4; Version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 161347 bases at least 040
Consensus quality: 162610 bases at least 030
Insert size: 163391; sum-of-ctrls

```


Insert size: 163373; 6.6% error; agarose-ff
Quality coverage: 5.19x in Q20 bases; sum-of-contigs quality
coverage: 5.21x in Q20 bases; agarose-ff

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 34236: contig of 34236 bp in length
* 34237 34336: gap of 100 bp
* 34337 47372: contig of 13036 bp in length
* 47373 47472: gap of 100 bp
* 47473 67025: contig of 19553 bp in length
* 67026 67125: gap of 100 bp
* 67126 71239: contig of 4114 bp in length
* 71240 71339: gap of 100 bp
* 71340 83903: contig of 12564 bp in length
* 83904 84003: gap of 100 bp
* 84004 91357: contig of 7354 bp in length
* 91358 91457: gap of 100 bp
* 91458 96452: contig of 4995 bp in length
* 96453 96552: gap of 100 bp
* 96553 148189: contig of 51637 bp in length
* 148190 148289: gap of 100 bp
* 148290 159778: contig of 11489 bp in length
* 159779 159878: gap of 100 bp
* 159879 164891: contig of 5013 bp in length.

FEATURES

Location/Qualifiers

misc_feature
1..164891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-444C12"
/clone_lib="RPC1-11.2"
1..34236
/note="assembly-fragment:00027
fragment_chain:1
clone_end:"7
vector_side:left"
34337..47372
/note="assembly-fragment:00312
fragment_chain:1"
47473..67025
/note="assembly-fragment:01923
fragment_chain:1"
67126..71239
/note="assembly-fragment:02094
fragment_chain:1"
71340..83903
/note="assembly-fragment:02213
fragment_chain:1"
84004..91357
/note="assembly-fragment:01292
fragment_chain:1"
91458..96452
/note="assembly-fragment:01391
fragment_chain:1"
96553..148189
/note="assembly-fragment:02045
fragment_chain:1"
148290..159778
/note="assembly-fragment:00427
fragment_chain:1"
159879..164891
/note="assembly-fragment:00773
fragment_chain:1
clone_end:sp6
vector_side:right"

misc_feature
48381 a 32179 c 32357 g 51067 t 907 others

BASE COUNT 48381 a 32179 c 32357 g 51067 t 907 others

ORIGIN

Query Match 63.8%; Score 115.4; DB 2: Length 164891;

Best Local Similarity 91.7%; Pred. No. 7.4e-20;
Matches 122; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

61 CAAAGATGACATTCATTAAGTACAGACACCTCTTGAAGAGATGAGTTGCTTCT 60
66468 CAAAGATGACATTCATTAAGTACAGACACCTCTTGAAGAGATGAGTTGCTTCT 66429

61 CCAATCCCGCCAGACGAGAGAGGCTTCTATTAATGTTGCAACAATGTTGATTAAGT 120
66468 CCAATCCCGCCAGACGAGAGAGGCTTCTATTAATGTTGCAACAATGTTGATTAAGT 66369

61 TGAATCTGTACA 133
66368 TGAATCTGTCTCA 66356

RESULT 7
AF095584
DEFINITION
Sus scrofa epithelial chloride channel protein (AEC) mRNA,
complete cds.
AF095584.1 GI:6002645

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000

AF095584 3079 bp mRNA MAM 13-OCT-2000


```

/codon_start=1
/protein_id="BAA37743.1"
/db_xref="gi:3721912"
/translation="MESLKSVPFLILLHLTGCVLSSESLIQLNNOGCEGVIAIDHDV
EDLEALLQHKMDVYQASPYLIEACBCKRKYFENVNAILIPESKAGPEYIRPLELEKKN
DAVLSTSPIDNDPEYRHYLIDACBCKRKYRILHPTDFLAGKLTQYQDDPRFFVEMHA
FRMVFENYNDDEKFEYLSKRGQQAQRCSAALTGNVRCGCGSGITNGKCVIDRWG
LYKNCVCFYNDPEYRHYLIDACBCKRKYFENVNAILIPESKAGPEYIRPLELEKKN

```

```

/gene="CAC2"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 2"
/protein_id="AA04839.1"
/db_xref="GI:576289"
/translation="MGLRGRGVFLVLCILHQSTSTIKLNNGFEDIVYIDSVPE
DEKTIIDEDVATVASTYLFEATKREFKFNKSTLIENKNNENQYRKREKNRHAD

```


VIVAPPLIPGDEPTTKOFTCEGKEVITHPTDILLCKKONETGPECKLVHEHAHL
RMGVFDEYNEDOPPEYRAKSKIEATRCVAGISGRNRYKCGGSLSRACRIDSTTKL
YKDCOEPPEDVOTERAKSIMFOSIDSVCEKNEHNOECSLONICNFRSTWEVI
SNSDEKNTIPMTVPPPPVSLIKRIRIYCLVLDKSGSMGKRLNRMOAARHL
LOTJENGSMWGMVHDSTATVKNLKIORSDEKRTLAAGLPTVGLGGTSCGKVA
FOVIGELHSOLDSSEVILLITGDENTASSCIDKVOASAIYHFIAGRAABEAVIENS
KITGSHFTVSDAONNGLIDAFGALISGNTDLSKSLQSLSKGILTSNMAMNTVI
IDSTVGKTFELITWNSLPSISLMDPSGTIMENTVATSKMALISIPGAKGVTA
LGNQAKAPETIESONGTEVLELDNGAGADSKNDGVSRFYAVENGRLSLKVAHG
YANVTAFIESONGTEVLELDNGAGADSKNDGVSRFYAVENGRLSLKVAHG
GANTARLKRPLNRRAAYIPGMVNGELIENAPREDEIDOTILEDPSRTASGAFV
VSQVSPUPPQPTPSQOTTDATVHEKIIITLTAPDNEDVGVORITIRISIL
DLRDSFDALQVNTDLSKEANSKESAPENISENATHIFELAIKSIDSKNTSK
VSNIAQVTLFIPQANPMDIDIPPTPPPKSHNSGVNISTVLVIGSVIVNPTLS
TTI"

BASE COUNT 1098 a 594 c 633 g 879 t

ORIGIN

Query Match 44.6%; Score 80.8; DB 9; Length 3204;
Best Local Similarity 65.6%; Pred. No. 6.6e-11;
Matches 118; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

OY 2 AAAAGTCACATTCATTAAGTAAAGAGCTCTATGAAAAAGATGCTGTTCTTC 61
DB 650 AGAGCATGAGATTCATTAAGTAAAGAGCTCTATGAAAAAGATGCTGTTCTTC 709
OY 62 CAATCCCGCAGAGGAGAGAGCTCTATTAATGTTTGACACATGTTGATTCATAGTT 121
DB 710 GATTAAGTACAAACAAAAAGAGCTCTATTAATGTTTGACACATGTTGATTCATAGTT 769
OY 122 GAATTCGTACAGAACAAACCAACAAAGAGCTCCAAAGAGCAAAATCAAAATGCG 181
DB 770 GAATTTGTACGAAAAAACCCATATATCAAGAGCTCCAAAGCTCAAAATCAAAATGCG 829

RESULT 10

AK000072 3221 bp mRNA PRI 22-FEB-2000
LOCUS Homo sapiens cDNA FLJ20065 f1s, clone COL01613, highly similar to
DEFINITION ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.

ACCESSION AK000072.1 GI:7019922
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_11b.COL clone:COL01613.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (sites)

REFERENCE Kawanabe, A., Hiki, J., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
TITLE Unpublished (2000)
JOURNAL 2 (bases 1 to 3221)
REFERENCE Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T.,
AUTHORS Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo, Department
of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES

location/Qualifiers
1..3221
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="COL01613"
/clone_11b="COL"
/tissue_type="colon"
/note="cloning vector pME18SFL3"
1..3221
/note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE
CHANNEL PROTEIN"
BASE COUNT 1105 a 600 c 634 g 882 t

Query Match 44.6%; Score 80.8; DB 9; Length 3221;
Best Local Similarity 65.6%; Pred. No. 6.6e-11;
Matches 118; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

OY 2 AAAAGTCACATTCATTAAGTAAAGAGCTCTATGAAAAAGATGCTGTTCTTC 61
DB 654 AGAGCATGAGATTCATTAAGTAAAGAGCTCTATGAAAAAGATGCTGTTCTTC 723
OY 62 CAATCCCGCAGAGGAGAGAGCTCTATTAATGTTTGACACATGTTGATTCATAGTT 121
DB 744 GATTAAGTACAAACAAAAAGAGCTCTATTAATGTTTGACACATGTTGATTCATAGTT 783
OY 122 GAATTCGTACAGAACAAACCAACAAAGAGCTCCAAAGAGCAAAATCAAAATGCG 181
DB 784 GAATTTGTACGAAAAAACCCATATATCAAGAGCTCCAAAGCTCAAAATCAAAATGCG 843

RESULT 11

AX092338 3265 bp DNA PAT 21-MAR-2001
LOCUS Sequence 69 from Patent WO0116318.
DEFINITION AX092338
ACCESSION AX092338
VERSION AX092338.1 GI:13444483

KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 3265)

REFERENCE Eaton, D.L., Filvaroff, E., Gerritsen, M.E., Goddard, A.,
AUTHORS Godowski, P.J., Grimaldi, C.D., Gurney, A.L., Watanabe, C.K. and
Wood, W.T.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
TITLE Patent: WO 0116318-A 69 08-MAR-2001;
JOURNAL Genentech, Inc. (US)

FEATURES
source location/Qualifiers
1..3265
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1159 a 596 c 632 g 878 t
ORIGIN

Query Match 44.6%; Score 80.8; DB 6; Length 3265;
Best Local Similarity 65.6%; Pred. No. 6.6e-11;
Matches 118; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

OY 2 AAAAGTCACATTCATTAAGTAAAGAGCTCTATGAAAAAGATGCTGTTCTTC 61
DB 646 AGAGCATGAGATTCATTAAGTAAAGAGCTCTATGAAAAAGATGCTGTTCTTC 705
OY 62 CAATCCCGCAGAGGAGAGAGCTCTATTAATGTTTGACACATGTTGATTCATAGTT 121
DB 706 GATTAAGTACAAACAAAAAGAGCTCTATTAATGTTTGACACATGTTGATTCATAGTT 765
OY 122 GAATTCGTACAGAACAAACCAACAAAGAGCTCCAAAGAGCAAAATCAAAATGCG 181
DB 766 GAATTTGTACGAAAAAACCCATATATCAAGAGCTCCAAAGCTCAAAATCAAAATGCG 825

RESULT	12	AF001264	1265 bp	mRNA	MAM	18-NOV-1997
LOCUS	AF001264					
DEFINITION	Bos taurus clone 4 endothelial adhesion molecule Lu-ECAM-1 mRNA, complete cds.					
ACCESSION	AF001264					
VERSION	AF001264.1					
KEYWORDS	GI:2623768					
SOURCE						
ORGANISM	cow.					
REFERENCE	Bos taurus					
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovine; Bos.					
TITLE	1 (bases 1 to 1265)					
JOURNAL	Elble, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R., Goodwin, A. and Pauli, B. U.					
REFERENCE	Cloning and characterization of Lu-ECAM-1 suggest it is an endothelial chloride channel					
AUTHORS	Unpublished					
TITLE	2 (bases 1 to 1265)					
JOURNAL	Elble, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R., Goodwin, A. and Pauli, B. U.					
REFERENCE	Direct Submission					
AUTHORS	Submitted (25-Apr-1997) Vet. Pathology, Cornell U., VMC C4-146, Ithaca, NY 14853, USA					
TITLE	Location/Qualifiers					
JOURNAL	1..1265					
REFERENCE	/organism="Bos taurus"					
AUTHORS	/db_xref="taxon:9913"					
TITLE	/tissue_type="lung"					
JOURNAL	/cell_type="aortic endothelial cells"					
REFERENCE	/clone="4"					
AUTHORS	109..1137					
TITLE	/note="endothelial adhesion molecule; chloride channel; secreted form"					
JOURNAL	/codon_start=1					
REFERENCE	/product="Lu-ECAM-1"					
AUTHORS	/protein_id="AA06532.1"					
TITLE	/db_xref="GI:2623769"					
JOURNAL	/translation="MVLCNVNLFLETLHLPGMKSMVNLINNGYDGLVAINSPDEPDEKLIENTKEVTAASYLFPHATKRRYFRVNSILIPWTKSGSEYFIPQESYDQADVIVANPYLKVGGDDPYTLOYGRCGKGYIHETPELILNNHIGSGRVRVHMAHLRMGFDEYNDVDPYIYSKRNTEIKSTHITGTVINVPFKCPGASCTISLDRSQGTGVLYEKCFLEPKRSQTAKESINEMFSLSHVEFCGTEHNTAEPMLOKMKCGKSTPQVIMNSVDYFQNTSPMTEBNPPTPTPSLKSQKVVCVILDKSGMSABEDIVLLALIKIRIKFKLIGNTI"					
BASE COUNT	440 a	231 c	231 g	363 t		
ORIGIN						
Query Match	41.0%; Score 74.2; DB 4; Length 1265;					
Best Local Similarity	64.0%; Pred. No. 3.4e-09;					
Matches 112; Conservative	0; Mismatches 63; Indels 0; Gaps 0;					
Y	7 ATGCATTCATTAATAAGTAAACAGAGCTATGCAAAAAGAGTGTAGTTGTTCTCCAAATC 66					
Db	738 ATGCAGAGCTGACTCACAGACAGCGGCTGTATGAACAAATGTACATTCCTCCAAAAAA 797					
Y	67 CCGCAGACGGGAGGCGTCTATATGTTGTCACAAACATGTTGATTCATATAGTGAATT 126					
Db	798 ATCCAGACTGCAAAAGGAAATCATATATGTTTATGCGAAGTCTCCATTCGTGACAGAAAT 857					
Y	127 CTGTACAGAACAAACCAACAAAGAAGCTCCAAACAAAGCAAAATCAAAAAATGC 181					
Db	858 TTGTACAGAAAAAACAACACAAATACAGAGCTCCAAACTACAAAAACAAATGTGC 912					
RESULT	13					
LOCUS	AF001263	2820 bp	mRNA	MAM	18-NOV-1997	
DEFINITION	Bos taurus clone 3 endothelial adhesion molecule Lu-ECAM-1 mRNA, complete cds.					
ACCESSION	AF001263					

```

VERSION      AF001263.1  GI:26237366
KEYWORDS
SOURCE
ORGANISM     Bos taurus
              Cow
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
              Bovidae; Bovinae; Bos.
REFERENCE     1 (bases 1 to 2820)
              Eble, R.C., Widom, J., U.
              Goodwin, A., and Pauli, B. U.
              Cloning and characterization of Lu-BCAM-1 suggest it is an
              endothelial chloride channel
              unpublished
              2 (bases 1 to 2820)
              Eble, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
              Goodwin, A., and Pauli, B. U.
              Direct Submission
              Submitted (25-APR-1997) Vet. Pathology, Cornell U., VMC C4-146,
              Ithaca, NY 14853, USA
FEATURES
              Location/Qualifiers
              1..2820
                /organism="Bos taurus"
                /db_xref="taxon:9913"
                /clone="3"
                /tissue_type="lung"
                /cell_type="aortic endothelial cells"
                194..2656
                /note="endothelial adhesion molecule; chloride channel"
                /codon_start=1
                /product="Lu-BCAM-1"
                /protein_id="AAB8653.1"
                /db_xref="GI:2623767"
                /translation="MVLCLANVLEFLTLHLPGKSSWVNLINNGYDVIIVAINSPVE
                DEKILNIKEMVTAETSLFPAFKRRVYFRNVSLIIPMTKSEYFIPOESDOAD
                VIVANPLKXDDPYLTLOYGCGEKGKIHTPEMLINNHAIYSGRVEVEMAHIL
                RMGLEDVNDNDPPYIRSKNTIEATRCSTHTGIVNYEKKCPGSGCTSLCRBSQV
                LYEKCTFLPKRSOTAKESIMFMSLSHSTVECTEKNHTDEAPLQNKCKGKSTWV
                IANSVDFQNTSPMTENMPPTPIPTSLSKQRYVCLVDSGMSADRLQNMQAEE
                LYLQVLEKSLVQWTFPDSVAELIQNHILRTTDNVYOKITAKLPVANGSTISICGL
                KAGQALIHSDQSTGSEIILLTDGEENELNSCEDVRSQALHTLALGPSAKELLE
                TLSMTGGYREFANKDLITGNAPSRISREGSITQOALQLESKALKITGKRVNGTV
                PVDSVNDPEFVYWTLOKPEYLQPKKKYKTSPEKDKLIRSRALQIPAIAT
                GTVYSLINNHASSQMLVYVTTTAKSPITPEVVAIAHMSOHTAHYSPMIVAQVSO
                GFLPEVLISVAIIETEDGHQVTLTENDNGAGRDYVNDGISYFYDYGNGYSLK
                VHAQRNNTALNRQPNKVLVPGVGVENKILINPREVKDKDLAKARIEFSRLTL
                SGGFYVSCAAPPNHSPVSPSKYITDLEAFKBDYIQLSWTAPGNVLDKKGASGSP
                MSRSHOVAKYLEIOLQHSQF"
BASE COUNT   940 a      565 c      560 g      755 t
ORIGIN
Query Match      41.0%; Score 74.2; DB 4; Length 2820;
Best Local Similarity 64.0%; Pred. No. 3,2e-09;
Matches 112; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
CY      7 ATGCACATTCATAAAGTACAGAGCCTATGAGAAAAGATGTGAGTTGTTCTCCAAATC 66
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
DB      823 ATGCAGAGTGACTCACAGACAGGCGCTGTATGACCAACCAAAATGTACATCTCTCCAAAAA 882
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
CY      67 CCGCAGACGAGGAGAGGCTCTATATGTTTGACACACATGTTGATTCTATATAGTTGAATT 126
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
EB      633 ATCCAGAGCTGCMAAGGATCCATATATGTTATGCAAGTCTCCATCTGTGACAGCAATT 942
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
CY      137 CTGTGACAAACAAAAACGACAAACAAAGAGCTCCAAACGAACGAAATCAAAAATGCG 181
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
EB      943 TTGTACAAAAAACACACATACAGAGGCTCCAAACCTACCAACCAACAAAATGTGCG 997
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
RESULT  14
AF001262      3288 bp      mRNA      MAM      18-NOV-1997
DEFINITION   Bos taurus gene 2 endothelial adhesion molecule Lu-BCAM-1 mRNA,
complete cds.

```


ACCESSION AF001262 GI:2623764
VERSION AF001261.1
KEYWORDS
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 3288)
AUTHORS Elble, R.C., Widom, J., Gruber, A.D., Abdel-Chany, M., Levine, R., Goodwin, A., and Pauli, B.U.
TITLE Cloning and characterization of Lu-ECAM-1 suggest it is an endothelial chloride channel
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3288)
AUTHORS Elble, R.C., Widom, J., Gruber, A.D., Abdel-Chany, M., Levine, R., Goodwin, A., and Pauli, B.U.
TITLE Direct Submission
JOURNAL Submitted (25-APR-1997) Vet. Pathology, Cornell U., VMC C4-146, Ithaca, NY 14853, USA
FEATURES
source location/Qualifiers
1..3288
/organism="Bos taurus"
/db_xref="taxon:9913"
/cell_type="aortic endothelial cells"
/clone="2"
/tissue_type="Lung"
59..2443
/note="endothelial adhesion molecule; chloride channel"
/codon_start=1
/product="Lu-ECAM-1"
/protein_id="AAB86530.1"
/db_xref="GI:2623765"
/translation="MWICLVNLTFLTHLPGKMSNVNLINNGYGIYAIINPSYDEKLENTKEMTEASTYLPHTKRRVFRNVSILPMTKSKSEPTFKOSTDQADVIVANPYLKGDDPYTLOYGRGEGKGIHFPNPLTNFNHLYSGRGVFEHMAHLRMGIDEPYVNDQPFYISRNKTLEATRCSTHIGINVEFKKPGSGCITSLCRDSQGLYEACTELPKRSQTAKESEIMPELSLSTKORVYCLVLDKSGMSAEDLPOKNOAEINNSVDFONTSPTEENPPTHPFSLSKORVYCLVLDKSGMSAEDLPOKNOAEIYLIOYIEKSLVGVTPDSVAEIONHNLRTIDNDNYOKITAKLPQVANGSTICGLKAGFOAIHSDSTSGSEITLLTDEGDEINSCFEVKKRSGAIHTHTALGPSAKLETLNMTGGIRFRANKDITGLTNAFSRISRSSTIQAOLQLESKALKITGRKRVNGTVPDSTVGNDFEVVVTWQPEVLVADPPGKRYKTSDFEKDLNHSALQIPLGNETGTWYSILNNHNSOMLTVTYVTRASPPIPIVIAAHNSOHTAHPSPMIYAYQSQGEPLVAGISVIAIETEDGHOVLELMDNGAGRDYKNGDISRYETDYGNRGYSLKVHAQARNNTARLNLROPNKVIYVPGYVNGKILNPPPEVYKDLAKKIDFESLTSGGSFTVSGAPPPGNHPSVFPSPSKITDLEAKRKEDYIQLSWTRPGNVLDKGES"
BASE COUNT 1043 a 707 c 645 g 893 t
ORIGIN

Query Match 41.0%; Score 74.2; DB 4; Length 3288;
Best Local Similarity 64.0%; Pred. No. 3.2e-09;
Matches 112; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 7 ATGCACATTCATTAAGTACAGAGCTCTATGAAAAAGATGTGATTTGTCCTCCATC 66
Db 688 ATGCACATTCATTCAGACAGAGCTCTATGAAAAAGATGTGATTTGTCCTCCATC 66
QY 67 CCGCCAGAGGAGAGGCTTCTATATGTTTGCACACATGTTGATTTCTAGTTGATT 126
Db 748 ATCCACAGAGCTGCAAGAGATCCATTTATGTTATGCCAAGCTCCATCTGTCGATTA 807
QY 127 CTGTACAGAACAAACACAAAGAGCTCCAAACAGCAAAATCAAAAATGC 181
Db 808 TTGTACAGAAAACACACATACAGAGCTCCAAACCTTACAAAACAAATGTGC 862

RESULT 15
AF001261 3317 bp mRNA MAM 18-NOV-1997
LOCUS AF001261
DEFINITION Bos taurus clone 1 endothelial adhesion molecule Lu-ECAM-1 mRNA, complete cds.

ACCESSION AF001261 GI:2623762
VERSION AF001261.1
KEYWORDS
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 3317)
AUTHORS Elble, R.C., Widom, J., Gruber, A.D., Abdel-Chany, M., Levine, R., Goodwin, A., and Pauli, B.U.
TITLE Cloning and characterization of Lu-ECAM-1 suggest it is an endothelial chloride channel
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3317)
AUTHORS Elble, R.C., Widom, J., Gruber, A.D., Abdel-Chany, M., Levine, R., Goodwin, A., and Pauli, B.U.
TITLE Direct Submission
JOURNAL Submitted (25-APR-1997) Vet. Pathology, Cornell U., VMC C4-146, Ithaca, NY 14853, USA
FEATURES
source location/Qualifiers
1..3317
/organism="Bos taurus"
/db_xref="taxon:9913"
/cell_type="aortic endothelial cells"
/clone="1"
/tissue_type="Lung"
63..2780
/note="endothelial adhesion molecule; chloride channel"
/codon_start=1
/product="Lu-ECAM-1"
/protein_id="AAB86529.1"
/db_xref="GI:2623763"
/translation="MWICLVNLTFLTHLPGKMSNVNLINNGYGIYAIINPSYDEKLENTKEMTEASTYLPHTKRRVFRNVSILPMTKSKSEPTFKOSTDQADVIVANPYLKGDDPYTLOYGRGEGKGIHFPNPLTNFNHLYSGRGVFEHMAHLRMGIDEPYVNDQPFYISRNKTLEATRCSTHIGINVEFKKPGSGCITSLCRDSQGLYEACTELPKRSQTAKESEIMPELSLSTKORVYCLVLDKSGMSAEDLPOKNOAEINNSVDFONTSPTEENPPTHPFSLSKORVYCLVLDKSGMSAEDLPOKNOAEIYLIOYIEKSLVGVTPDSVAEIONHNLRTIDNDNYOKITAKLPQVANGSTICGLKAGFOAIHSDSTSGSEITLLTDEGDEINSCFEVKKRSGAIHTHTALGPSAKLETLNMTGGIRFRANKDITGLTNAFSRISRSSTIQAOLQLESKALKITGRKRVNGTVPDSTVGNDFEVVVTWQPEVLVADPPGKRYKTSDFEKDLNHSALQIPLGNETGTWYSILNNHNSOMLTVTYVTRASPPIPIVIAAHNSOHTAHPSPMIYAYQSQGEPLVAGISVIAIETEDGHOVLELMDNGAGRDYKNGDISRYETDYGNRGYSLKVHAQARNNTARLNLROPNKVIYVPGYVNGKILNPPPEVYKDLAKKIDFESLTSGGSFTVSGAPPPGNHPSVFPSPSKITDLEAKRKEDYIQLSWTRPGNVLDKGRANSYIIRKSTFMDQEDFDNATLVNTSLNLKESGSENFEPKPEHREVENTGYFLYSQAINFANLISEVSHIYQAIKFIPLPDSVHDLGTKISEITLAILGLPMIPSV"
BASE COUNT 1084 a 676 c 634 g 923 t
ORIGIN

Query Match 41.0%; Score 74.2; DB 4; Length 3317;
Best Local Similarity 64.0%; Pred. No. 3.2e-09;
Matches 112; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 7 ATGCACATTCATTAAGTACAGAGCTCTATGAAAAAGATGTGATTTGTCCTCCATC 66
Db 692 ATGCACATTCATTCAGACAGAGCTCTATGAAAAAGATGTGATTTGTCCTCCATC 66
QY 67 CCGCCAGAGGAGAGGCTTCTATATGTTTGCACACATGTTGATTTCTAGTTGATT 126
Db 752 ATCCACAGAGCTGCAAGAGATCCATTTATGTTATGCCAAGCTCCATCTGTCGATTA 811
QY 127 CTGTACAGAACAAACACAAAGAGCTCCAAACAGCAAAATCAAAAATGC 181
Db 812 TTGTACAGAAAACACACATACAGAGCTCCAAACCTTACAAAACAAATGTGC 866

Search completed: April 4, 2002, 02:57:29
Job time: 60131 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 03:17:23 ; Search time 1321.64 Seconds
(without alignments)
142.710 Million cell updates/sec

Title: US-09-049-696-5

Perfect score: 220
Sequence: 1 CTTAGTGAATTCGTGACA.....GACAAATCTGGAACATGGC 220

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_1101.*
1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	220	100.0	2742	22	AAH46102 Human CLCA1 coding
2	220	100.0	2745	22	AAF81927 Human ICAC-1 nucl
3	220	100.0	2825	22	AAH46124 Human ICAC1 CDNA,
4	220	100.0	2854	22	AAH34879 Human colon cancer
5	220	100.0	2854	22	AAH81787 Human secreted pro
6	220	100.0	2867	22	AAH33285 Human colon cancer
7	220	100.0	3109	22	AAH35019 Human colon cancer
8	220	100.0	3111	22	AAZ09840 Human membrane spa
9	220	100.0	3311	22	AAI29502 C902P determined c
10	162	73.6	2739	22	AAH46101 Mouse Gob-5 coding
11	162	73.6	2843	22	AAH46120 Mouse Gob-5 CDNA,

12	162	73.6	2931	20	AAF81925 Murine ICAC-1 nuc
13	132.6	60.3	3216	21	AAA64335 Clone 2516888 of a
14	132.6	60.3	3265	21	AAZ65095 Membrane-bound pro
15	132.6	60.3	3265	22	AAF92092 Human PRO1124 cDNA
16	132.6	60.3	3265	22	AAF44241 Human PRO1124 (DNQ)
17	82.4	37.5	2454	22	AAH82881 Human tumor-associ
18	80.8	36.7	2784	21	AAZ24658 Human lung cancer-
19	80.8	36.7	2784	21	AAZ65897 Human lung cancer-
20	80.8	36.7	3362	20	AAZ24657 Human lung cancer-
21	80.8	36.7	3362	21	AAZ65896 Human lung cancer-
22	80.8	36.7	3351	20	AAZ24653 Human lung tumor a
23	80.8	36.7	3951	21	AAH65892 Human lung cancer-
24	80.8	36.7	8031	21	AAH65950 Human lung cancer-
25	76	34.5	3190	20	AAF81926 Human ICAC-2 nucl
26	74.2	33.7	3156	21	AAZ51625 Human membrane cha
27	49.2	22.4	307	22	AAH72076 Human cervical can
28	49.2	22.4	333	22	AAH71005 Human cervical can
29	49.2	22.4	423	22	AAH69200 Human cervical can
30	32.6	14.8	698	18	AAV06140 Viral infection ge
31	32.4	14.7	14460	21	AAZ93815 Olfactory receptor
32	32.2	14.6	1814	22	AAV74769 Staphylococcus aur
33	32.2	14.6	2434	22	AAH57479 Human SAR2 nucleot
34	31.4	14.3	414	22	AAI36754 Probe #5440 used t
35	31.4	14.3	910715	20	AAH20248 Borrelia burgdorfe
36	31	14.1	524	16	AAH35109 Down-regulated sen
37	31	14.1	1478	21	AAZ46267 Arabidopsis thalia
38	31	14.1	1480	21	AAH35460 Human secreted pro
39	30.8	14.0	321	21	AAH47760 Zea mays DNA fragm
40	30.8	13.9	1009	21	AAH89888 Human bone marrow
41	30.6	13.9	1945	22	AAH89888 Oligonucleotide D1
42	30.4	13.8	936	22	AAH89888 Oligonucleotide D1
43	30.4	13.8	936	22	AAH89888 Oligonucleotide D1
44	30.4	13.8	936	22	AAH89888 Oligonucleotide D1
45	30.4	13.8	936	22	AAH89888 Oligonucleotide D2

ALIGNMENTS

AAH46102	1	AAH46102 standard; DNA; 2742 bp.
AAH46102	1	AAH46102
AC	AAH46102	
XX	11-SEP-2001	(first entry)
DE	Human CLCA1 coding sequence; SEQ ID NO:4.	
XX	Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;	
KW	expression inhibition; antisense therapy; gene therapy;	
KW	chronic obstructive pulmonary disease; bronchial asthma; antilasthmatic;	
KW	ds.	
OS	Homo sapiens;	
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..2742
FT		/*tag= a
FT		/product= "Human CLCA1"
FT		/note= "No stop codon given in the specification"
XX	MO200138530-A1.	
PN	31-MAY-2001.	
XX		
PD	22-NOV-2000; 2000WO-JP08232.	
XX		
PF	24-NOV-1999; 99UP-033479.	
XX		
FR	27-APR-2000; 2000UP-0127589.	
PR		
XX		
PA	(TAKE) TAKEDA CHEM IND LTD.	

XX Nakanishi A, Morita S;
 PI WPI: 2001-355935/37.
 DR P-PSDB; AAB73716.
 XX
 PT New antisense nucleotide, useful for treatment and prevention of
 PT bronchial asthma and chronic obstructive pulmonary disease
 XX
 PS Claim 3; Page 82-84; 104pp; Japanese.
 XX
 CC The invention relates to an antisense nucleotide targeted to the mouse
 CC Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
 CC the CLCA1 gene (coding sequence shown in AAH46102). The invention also
 CC relates to an antibody specific for the Gob-5 protein, medical and
 CC diagnostic compositions containing the antisense nucleotide or the
 CC antibody, and methods and kits for screening for compounds which inhibit
 CC the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
 CC The antisense oligonucleotides and antibody are therefore useful for the
 CC treatment and prevention of bronchial asthma and chronic obstructive
 CC pulmonary disease. The present sequence represents the human CLCA1
 CC gene coding sequence.
 XX
 SO Sequence 2742 BP; 833 A; 616 C; 623 G; 670 T; 0 other;

Query Match 100.0%; Score 220; DB 22; Length 2742;
 Best Local Similarity 100.0%; Pred. No. 3.9e-56;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTATAGTTGAATTCCTGTCAGACAAACCAACAAAGAGCTCCAAACAGCAAAATC 60
 Db |||||||
 Db 734 ctatagttgaattctctgtacagaaacacacaaagaagctccaaagcaaatc 793
 QY 61 AAAAATGCAATCTCCGAGACATGGAAGTATCCGTATTCGAGGACTTTAAGAAAA 120
 Db |||||||
 Db 794 aaaaatgcaatctccgagacacatggaagtatccgtatctcgaggaactttaagaaa 853
 QY 121 CCACCTCTATGACAAACAGCACCACCAATCCACCTTCTCATTTGCTGAGATTGGACAAA 180
 Db |||||||
 Db 854 ccacctctatgacaacacagaccacaaatccacctctcatctgtaggattgacaaa 913
 QY 181 GAATTGCTGTTAGTCTTACAAATCTGAGAGCATGCGC 220
 Db |||||||
 Db 914 gaattgctgtttagtctcttgacaacatctggaagcatgyc 953

RESULT 2
 ID AAF81927 standard; cDNA; 2745 BP.
 AC AAF81927;
 XX
 DT 13-JUN-2001 (first entry)
 XX
 DE Human ICACC-1 nucleotide sequence.
 XX
 KW ICACC-1; ICACC-2; asthma; atopic allergy; asthma-related disorder;
 KW interleukin 9 induced calcium activated chloride channel; IL-9;
 KW calcium activated chloride channel; anti-allergic; anti-asthmatic;
 KW anti-inflammatory; immunomodulatory; cystic fibrosis;
 KW inflammatory bowel disease; autoimmune disease; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key 1. 2745 Location/Qualifiers
 FT CDS
 FT /*tag= a
 FT /product= "ICACC-1"
 FT /note= "IL-9 induced calcium activated chloride channel"
 XX
 XX WO944620-A1.

PD 10-SEP-1999.
 XX
 XX 03-MAR-1999; 99WO-US04703.
 XX
 ER 03-MAR-1998; 98US-0076815.
 XX
 PA (MGA-1) MAGAININ PHARM INC.
 XX
 PI Holroyd RJ, Levitt RC, Maloy WL, Louahed J, McLane M;
 PI Nicolides NC, Zhou Y, Dong Q;
 XX
 DR WPI: 1999-550979/46.
 DR P-PSDB; AAB74824.
 XX
 PT New nucleic acid encoding calcium activated chloride channel, used to
 PT identify, e.g. specific modulators for treating atopic allergy
 XX
 PS Claim 1; Fig 4B; 75pp; English.
 XX
 CC The present sequence encodes the human interleukin 9 (IL-9) induced
 CC calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins
 CC have anti-allergic, anti-asthmatic, anti-inflammatory and
 CC immunomodulatory activities. Compounds (A) that downregulate ICACC are
 CC used to alleviate asthma (or more generally atopic allergy), while those
 CC (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and
 CC inflammatory bowel disease (IBD) (or other autoimmune diseases).
 CC Measuring levels of ICACC-1 can be used in the diagnosis of asthma
 CC (increased levels) or IBD (reduced levels), also for monitoring
 CC treatment of these conditions. The ICACC proteins can be used:
 CC (1) to raise specific antibodies (Ab), useful: (a) as immunoassay
 CC reagents, and (b) as therapeutic (A); (11) as fragments or mutants, as
 CC (A) to disrupt bonding between ICACC and its ligands; and (11) to
 CC identify modulators and binding partners. ICACC polynucleotides can be
 CC used to generate transgenic animals or recombinant cells, used to screen
 CC for antagonists, also as a source of therapeutic antisense agents or
 CC diagnostic probes (for quantifying mRNA expression, e.g. for
 CC identification of modulators).
 XX
 SO Sequence 2745 BP; 833 A; 616 C; 625 G; 671 T; 0 other;

Query Match 100.0%; Score 220; DB 20; Length 2745;
 Best Local Similarity 100.0%; Pred. No. 3.9e-56;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTATAGTTGAATTCCTGTCAGACAAACCAACAAAGAGCTCCAAACAGCAAAATC 60
 Db |||||||
 Db 734 ctatagttgaattctctgtacagaaacacacaaagaagctccaaagcaaatc 793
 QY 61 AAAAATGCAATCTCCGAGACATGGAAGTATCCGTATTCGAGGACTTTAAGAAAA 120
 Db |||||||
 Db 794 aaaaatgcaatctccgagacacatggaagtatccgtatctcgaggaactttaagaaa 853
 QY 121 CCACCTCTATGACAAACAGCACCACCAATCCACCTTCTCATTTGCTGAGATTGGACAAA 180
 Db |||||||
 Db 854 ccacctctatgacaacacagaccacaaatccacctctcatctgtaggattgacaaa 913
 QY 181 GAATTGCTGTTAGTCTTACAAATCTGAGAGCATGCGC 220
 Db |||||||
 Db 914 gaattgctgtttagtctcttgacaacatctggaagcatgyc 953

RESULT 3
 ID AAH46124 standard; cDNA; 2825 BP.
 AC AAH46124;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Human CLCA1 cDNA, SEQ ID NO:26.
 XX
 XX Human CLCA1; goblet cell; mouse Gob-5 orthologue; drug screening;

KM	expression inhibition; antisense therapy; gene therapy;
RN	chronic obstructive pulmonary disease; bronchial asthma; antiasthmatic;
KW	ss.
XX	Homo sapiens.
OS	
FH	Location/Qualifiers
FT	23..2767
FT	/tag= a
FT	/product= "Human CLCA1"
FT	/transcript= (pos:476..478, aa:Lys)
PN	WO200138530-A1.
XX	
PD	31-MAY-2001.
XX	
PE	22-NOV-2000; 2000WO-JP08232.
XX	
PR	24-NOV-1999; 99JP-0333479.
PR	27-APR-2000; 2000JP-0127589.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
XX	
PI	Nakanishi A, Morita S:
DR	WPI; 2001-355935/37.
DR	P-PDB; AAB73716.
PT	New antisense nucleotide, useful for treatment and prevention of
PT	bronchial asthma and chronic obstructive pulmonary disease -
PS	Example 5; Page 92-94; 104pp; Japanese.
XX	
CC	The invention relates to an antisense nucleotide targeted to the mouse
CC	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC	the CLCA1 gene (coding sequence shown in AAH46102). The invention also
CC	relates to an antibody specific for the Gob-5 protein, medical and
CC	diagnostic compositions containing the antisense nucleotide or the
CC	antibody, and methods and kits for screening for compounds which inhibit
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
CC	The antisense oligonucleotides and antibody are therefore useful for the
CC	treatment and prevention of bronchial asthma and chronic obstructive
CC	pulmonary disease. The present sequence represents human CLCA1 cDNA.
SQ	Sequence 2825 BP; 864 A; 626 C; 637 G; 698 T; 0 other;
Query Match	100.0%; Score 220; DB 22; Length 2825;
Best Local Similarity	100.0%; Pred. No. 4e-56;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	1 CTATGTTGAATTCTGTGCAGAAACAAACACACAACAAAGACGTCCTCAAACAGAACAATC 60
Dd	756 ctatggtgaattctgtgcagaaacaaaaccacaagaagcctccaaacaagcaaatc 815
OY	61 AAAAATGCAATCTCCGAAGCACATGGGAGTATCGTGATTTCTGAGACTTTAAGAAAA 120
Dd	816 aaaatgcaatctccgaagcacatgggagtatcggtgatcttgagactttaagaaaa 875
OY	121 CCATGCCCATATGACAAACAGCCACCATAATCCCACTTCTCATTTGCTGCAGATTGGACAAA 180
Dd	876 ccatagcccatatgacaaacagccaccataatccaccttcctcatgtcagattgagacaaa 935
OY	181 GAATGTGCTGTTTAGTCCCTTGACAATCTGGAAGCATGGC 220
Dd	936 gaatgtgttgtttagtcccttgacaatctggaagcatggc 975
RESULT 4	
ID	AAH34879 standard; cDNA; 2854 BP.
XX	
XC	AAH34879;

```

XX 03-SEP-2001 (first entry)
DE Human colon cancer antigen encoding cDNA SEQ ID NO:1961.
XX Human colon cancer; colon cancer antigen; diagnosis; detection;
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; chromosome 1; ss.
XX Homo sapiens.
XX WC0300122920-A2.
XX
XX PD 05-APR-2001.
XX PF 28-SEP-2000; 2000WO-US26524.
XX PR 29-SEP-1999; 99US-0157137.
XX PR 03-NOV-1999; 99US-0163280.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
XX MP: 2001-235357/24.
XX P-ESDB: AAG75474.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 1; Page 3462-3463; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytoskeletal activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing P.
XX inactive proteins or to supplement the patients own production of P.
XX Additionally, N may be used to produce the colon cancer-associated Ps,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAG77789 represent sequences used in the exemplification of the
XX present invention.
XX
XX N.B. Pages 666 to 662 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;
XX
XX
XX Query Match 100.0%; Score 220; DB 22; Length 2854;
XX Best local Similarity 100.0%; Pred. No. 4e-56;
XX Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```


XX	AAAF81787	standard; cDNA; 2854 BP.
XX	AAAF81787;	
DT	12-JUN-2001	(first entry)
XX	Human secreted protein gene 1 SEQ ID NO:11.	
XX	Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;	
KM	dermatological; immunosuppressive; antiinflammatory; anti-HIV;	
KM	immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;	
KM	ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;	
KM	antialzheimer; antiparkinsonian; antitumor; antitumor; antitumor;	
KM	immune disorder; hyperproliferative disorder; cardiovascular disease;	
KM	cancer; angiogenic disorder; neurological disorder; infectious disease;	
KM	wound healing; regeneration; chemotaxis; chromosome 1; ss.	
OS	Homo sapiens.	
PN	WO200112775-A2.	
XX	22-FEB-2001.	
PD	16-AUG-2000; 2000WO-US22325.	
PF	17-AUG-1999; 99US-0149182.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Ni J, Florence KA, Fiscella M, Wei P, Baker KP;	
PI	Bitse CE, Young PE, Komatsoulis GA, Moore PA, Soppet DR;	
XX	WPI; 2001-147550/15.	
XX	P-PSDB; AAF81787.	
DR	Nucleic acids encoding 25 human secreted polypeptides, useful for	
PT	preventing, diagnosing and/or treating e.g. cancers, Parkinson's	
PT	disease and diabetic retinopathy -	
XX	Claim 1; Page 441; 485pp; English.	
PS	AAAF81787 to AAF81817 encode the human secreted proteins given in AAF81787	
XX	to AAF81772. Human secreted proteins can have activities based on the	
CC	tissues and cells they are expressed in. Example of activities include:	
CC	immunomodulatory; antisclerotic; dermatological; immunosuppressive;	
CC	antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;	
CC	vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;	
CC	anticonvulsant; antialzheimer; antiparkinsonian; antitumor; and	
CC	vulnary. Human secreted proteins can be used in gene therapy and	
CC	vaccine. Human secreted protein nucleotide sequences (NAM1) and proteins	
CC	(PEP1) may be used in the prevention, diagnosis and treatment of diseases	
CC	associated with inappropriate polypeptide expression. For example, NAM1	
CC	and PEP1 may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of proteins by expressing inactive proteins or	
CC	to supplement the patient's own production of polypeptides. Disorders that	
CC	may be prevented, diagnosed and/or treated include immune disorders,	
CC	hyperproliferative disorders (e.g. cancers), cardiovascular diseases,	
CC	angiogenic disorders, neurological disorders, infectious diseases and/or	
CC	for promoting wound healing, regeneration and/or chemotaxis. AAF81778 to	
CC	AAAF81786 and AAF81787 represent sequences used in the exemplification of	
CC	the present invention.	
XX	Sequence 2854 BP; 885 A; 628 C; 643 G; 698 T; 0 other;	

Query Match	100.0%;	Score 220;	DB 22;	Length 2854;
Best Local Similarity	100.0%;	Pred. No. 4e-56;		
Matches 220;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	CTATGTTGAATTCGTGTACAGACAAAAACCCAAACAAAGAAAGCTCCAAACAGCAAAATC	60
Dp	768	ctatggttgaaattctgtacagaacaaacccacacaaagaagctcccaacaagaacaaatc	827
QY	61	AAAAATGCATCTCCGAGACGACATGGGGAAGATCCGTGATTCAGAGACTTTAAGAAAA	120
Dp	828	aaaaatgaatctccgaagacacatggaagatgcgtgattctcgagacttaagaana	887
QY	121	GCACCTCATATGACACACAGCCACCACCAATCCCACTTCTATTGCTGCAATTGGACAA	180
Dp	888	ccacctcctatgacaacaacagccaccaaatcccaactctcatgtcgtcagattggacaaa	947
QY	181	GAATTGTGTTAGTCCCTTGACAAATCTGCAAGCATGGC	220
Dp	948	gaattgtgtttagtcccttgacaaatctggaagcatggc	987

RESULT 6
 AAH33285
 ID AAH33285 standard; CDNA; 2867 BP.
 XX
 AC AAH33285;
 XX
 DT 03-SEP-2001 ((first entry))
 XX
 DE Human colon cancer antigen encoding cDNA SEQ ID NO:341.
 XX
 EN Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 1; ss.
 XX
 OS Homo sapiens.
 XX
 PN WC200122920-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000MO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (PUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR WPI; 2001-235357/24.
 XX
 F P-PSDB; AAG73854.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 XX
 FT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 Claim 1; Page 2452-2453; 9803pp; English.
 XX
 PS AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAG77789 represent sequences used in the exemplification of the
 CC present invention.
 CC
 CC K.R. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

PR	27-APR-2000; 2000JP-0127589.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
XX	
PI	Nakanishi A, Morita S;
XX	
DR	WPI; 2001-355935/37.
XX	P-PsDB: AAB73715.
PT	New antisense nucleotide, useful for treatment and prevention of
XX	bronchial asthma and chronic obstructive pulmonary disease
PS	Claim 3; Page 80-82; 104pp; Japanese.
XX	
CC	The invention relates to an antisense nucleotide targeted to the mouse
XX	Gob-5 gene (coding sequence shown in AAH46101) or its human counterpart,
CC	the CLCA1 gene (coding sequence shown in AAH46102). The invention also
XX	relates to an antibody specific for the Gob-5 protein, medical and
CC	diagnostic compositions containing the antisense nucleotide or the
XX	antibody, and methods and kits for screening for compounds which inhibit
CC	the protein. Gob-5 and CLCA1 are proteins expressed by goblet cells.
XX	The antisense oligonucleotides and antibody are therefore useful for the
CC	treatment and prevention of bronchial asthma and chronic obstructive
XX	pulmonary disease. The present sequence represents the mouse Gob-5
CC	gene coding sequence.
XX	
SQ	Sequence 2739 BP; 784 A; 687 C; 655 G; 613 T; 0 other;
XX	
Query Match	73.6%; Score 162; DB 22; Length 2739;
Best Local Similarity	83.9%; Pred.No.7.9e-39;
Matches 183; Conservative	0; Mismatches 35; Indels 0; Gaps 0
OY	1 CTATAGTTGAATTCTGTGTCAGACAACAAACCACACAAGAAGACTCCAAACAAGCAAAATC 60
Db	
	737 ctgtggttgaaatctcgtacagaaaaatacaaatcaagaaggccccaattgacaaaacc 796
OY	61 AAAAATGCATCTCCGAAGCACATGGGAAGTATCCGTGATTCTGAGACTTTAAGAAAA 120
Db	
	797 aacgatgcgaatctccgaagcacgtggaagtcatcogaagatctgagacttcaagcaaa 856
OY	121 CCATCTCGTAGCAGACACAGCCACCACCAATCCCACTTCTCATTCGTGCGAGATTGGACAA 180
Db	
	857 ccactcccatcgacagcccagccacctgcagccacccttcactctcgtgcaaatltygacaaa 916
OY	181 GAATTGTGTATTAGTCCGTGACAATAATCGAAGCATG 218
Db	
	917 gaattgttgttagttcttgataaagtcggagacatg 954
RESULT 11	
ID	AAH46120
XX	AAH46120 standard; cDNA; 2843 BP.
XX	
AC	AAH46120:
XX	
D7	11-SEP-2001 (first entry)
XX	
DE	Mouse Gob-5 cDNA, SEQ ID NO:22.
XX	
KW	Mouse Gob-5; murine; goblet cell; human CLCA1 orthologue; drug screening;
XX	expression inhibition; antisense therapy; gene therapy; bronchial asthma;
RN	chronic obstructive pulmonary disease; antiasthmatic; ss.
OS	Mus sp.
XX	
FH	Key Location/Qualifiers
FT	CDS 15..2756
XX	/tag= a
FT	/product= "Mouse Gob-5"
XX	
RN	WO200138530-A1.
XX	

[illegible]

FT	CDS		B..2749	/tag= "a /product= "ICACC-1" /note= "IL-9 induced calcium activated chloride channel"
PD	W09944620-AI.			
PX	10-SEP-1999.			
PE	03-MAR-1999; .	99MO-US04703.		
PR	03-MAR-1998;	98US-0076815.		
PA	(MAGA-) MAGAININ PHARM INC.			
PI	Holroyd KJ, Levitt RC, Maloy WL,	Louahed J, McLane M:		
PT	Nicolaides NC, Zhou Y, Dong Q;			
PS	New nucleic acid encoding calcium activated chloride channel, used to identify, e.g. specific modulators for treating atopic allergy - Claim 2; Fig 2; 75pp; English.			
CC	The present sequence encodes the murine interleukin 9 (IL-9) induced calcium activated chloride channel 1 (ICACC-1) protein. ICACC proteins have anti-allergic, anti-asthmatic, anti-inflammatory and immunomodulatory activities. Compounds (A) that downregulate ICACC are used to alleviate asthma (or more generally atopic allergy), while those (B) that upregulate ICACC are used to treat cystic fibrosis (CF) and inflammatory bowel disease (IBD) (or other autoimmune diseases). Measuring levels of ICACC-1 can be used in the diagnosis of asthma (increased levels) or IBD (reduced levels), also for monitoring treatment of these conditions. The ICACC proteins can be used: (1) to raise specific antibodies (Ab), useful: (a) as immunoassay reagents, and (b) as therapeutic (A); (ii) as fragments or mutants, as (A) to disrupt bonding between ICACC and its ligands; and (iii) to identify modulators and binding partners. ICACC polynucleotides can be used to generate transgenic animals or recombinant cells, used to screen for antagonists, also as a source of therapeutic antisense agents or diagnostic probes (for quantifying mRNA expression, e.g. for identification of modulators).			
SQ	Sequence 2931 BP; 867 A; 719 C; 684 G; 661 T; 0 other;			
Query Match	73.6%;	Score 162;	DB 20;	Length 2931.
Best Local Similarity	83.9%;	Pred No. 8.1e-39;		
Matches 183;	Conservative 0;	Mismatches 35;	Indels 0;	Gaps 0;
OY	1 CTATAGTGAATTCTGTCACGAACAACCAACAAAGAACGTCCAACAAGCAAATC	60		
Db				
	744 ctgttgatgaattctgtaacagaataaacaccacaatcaagaaaggccccaattgcctaacc	803		
OY	61 AAAATGCATTTGCCGAAGCACATGGGAAGTGATCGGTGATCTTAGAGACTTTAAGAAA	120		
Db				
	804 aaagtataaatctcgaagcacgttggaaatcatcggaactctagagacctcaagcaaa	863		
OY	121 CCATCCCATATGACACACAGCCACCAAAATCCCACTTCGATTCGTAAGTGGACAA	180		
Db				
	864 ccaatccccatgccagcccacgaacctgaaccaaccacatcttcgaactgtcgaatatgtgacaaa	923		
OY	181 GAATTGTGTGTTAGCTTGACAAATCTGGAAGCATC	218		
Db				
	924 gaatttgtctagttctctgatgaatgcgcgaagcaty gatl			
RESULT	13			
AAB64335				
ID	AAB64335 standard; DNA; 2616 BP.			
XX				

AA04335;
20-DEC-2000 (first entry)
Clone 2516888 of a novel gene associated with colon disease.
Colon disease; colon cancer gene; carbonic anhydrase; galectin;
carcinoembryonic antigen; colorectal carcinoma tumor associated antigen;
fatty-acid binding protein; glutathione peroxidase; guanylin;
cytokeratin; cadherin; intestinal mucin; colon cancer; Crohn's disease;
metastatic bowel cancer; atrophic gastritis; cholecystitis;
irritable bowel syndrome; ulcerative colitis; ss.
Homo sapiens;
WO200050588-A2.
31-AUG-2000.
01-FEB-2000; 2000MO-US02595.
22-FEB-1999; 99US-0255381.
(INCY-) INCYTE PHARM INC.
Walker MG, Volkmuth W, Klingler TM, Lal P;
WPI; 2000-556397/51.
Novel gene associated with colon cancer and coexpressed with one or
more known colon cancer genes in a number of biological samples, for
use in gene therapy -
Claim 2; Page 30-31; 33pp; English.

The present sequence represents a novel gene which is associated with
colon disease. The gene is coexpressed with one or more colon cancer
genes selected from carbonic anhydrase I, II and IV (CA I, II and IV),
carcinoembryonic antigen family of proteins (ceas), colorectal carcinoma
tumor associated antigen (CO-029), down-regulated in adenoma (dra),
fatty-acid binding protein (fabp), galectin (galec), glutathione
peroxidase (gpx2), guanylin (guan), cytokeratin 8 and 20 (Ker 8 and 20),
cadherin (cadher) or intestinal mucin (muc-2). The present polynucleotide
sequence is useful for diagnosing a disease or condition associated
with altered expression of the gene. The polynucleotide and encoded
polypeptides are useful for treating or preventing such diseases. They
are useful for diagnosing colon cancer, metastatic colon cancer, atrophic
gastritis, cholecystitis, Crohn's disease, irritable bowel syndrome and
ulcerative colitis.

Sequence 2616 BP; 887 A; 488 C; 515 G; 726 T; 0 other;

Query Match 60.3%; Score 132.6; DB 21; Length 2616;
Best Local Similarity 75.3%; Pred. No. 4.6e-30;
Matches 165; Conservative 0; Mismatches 54; Indels 0; Gaps 0

1 CTATAGTGAATTCTGTTCAGACAACAAAACCAACAAAGAAGACTCCAAACGAAATC 60
194 ctgttgatgaatttgttaagaaaaaacccataacaaagaagctccaagctcaaaaaca 253
51 AAAAATGCATATCCGAAGACACATGAGATGATCGTGATTTCTGAGACTTAAGAAAA 120
254 taaggtagaatlttagaagtacatggagtggttttagcatctcgaggttttaaaaaa 313
121 CCAGTCGATAGAACACAGACGCCAACCAATCCCACTTCTCATGTGGCGAGATTGACAAA 180
314 ccaatacccattgtagaccaccaacctcccaacctgccttcattgcatggaatcaglcacaa 373
191 GAATGTGCTGTATGCTCTTGACAAATTCGAAAGCATG 219
374 gaattgtgtcgtactgtgataagcttggaagatgg 412

RESULT 14
AAZ65095
ID AAZ65095 standard; cDNA; 3265 BP.
XX
AC AAZ65095;
XX
DT 05-APR-2000 (first entry)
XX
DE Membrane-bound protein PRO1124 encoding cDN.
XX
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping; ss.
XX
OS Homo sapiens.
XX
PN WO963088-A2.
XX
PD 09-DEC-1999.
XX
PF 02-JUN-1999; 99WO-0512252.
XX
PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.

PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091544.
PR 01-JUL-1998; 98US-0091547.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 20-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097661.
PR 26-AUG-1998; 98US-0097951.

PR 26-AUG-1998; 98US-0097952.
 PR 26-AUG-1998; 98US-0097954.
 PR 26-AUG-1998; 98US-0097955.
 PR 26-AUG-1998; 98US-0097971.
 PR 26-AUG-1998; 98US-0097974.
 PR 26-AUG-1998; 98US-0097978.
 PR 26-AUG-1998; 98US-0097979.
 PR 26-AUG-1998; 98US-0097986.
 PR 26-AUG-1998; 98US-0098014.
 PR 31-AUG-1998; 98US-0098525.
 PR 16-SEP-1998; 98US-0100634.
 PR 12-JAN-1999; 99US-0115565.
 (GETH) GENENTECH INC.
 PI Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
 PI Wood WI, Yuan J;
 DR WPI: 2000-072883/06.
 DR P-PSDB: AAY66749.
 PT Membrane-bound proteins and related nucleotide sequences -
 PS Claim 2; Fig 273; 822pp; English.
 CC The invention provides membrane-bound PRO polypeptides and
 CC polynucleotides encoding them. The PRO sequences of the invention were
 CC identified based on extracellular domain homology screening. The PRO
 CC sequences have homology with proteins including LDL receptors, TIE
 CC ligands and various enzymes. The membrane-bound proteins and receptor
 CC molecules are useful as pharmaceutical and diagnostic agents. Receptor
 CC immunoadhesins, for instance, can be used as therapeutic agents to block
 CC receptor-ligand interactions. The membrane-bound proteins can also be
 CC employed for screening of potential peptide or small molecule inhibitors
 CC of the relevant receptor/ligand interaction. The PRO encoding sequences
 CC are useful as hybridization probes, in chromosome and gene mapping and in
 CC the generation of antisense RNA and DNA. PRO nucleic acid sequences
 CC will also be useful for the preparation of PRO polypeptides, especially
 CC by recombinant techniques.
 XX
 XX Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;
 Query Match 60.3%; Score 132.6; DB 21; Length 3265;
 Best Local Similarity 75.3%; Pred. No. 4.9e-30;
 Matches 165; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 1 CTATAGTTGATTCGTGACAGCAACAAAGAGAGCTCCAAACAGCAAGCAATC 60
 Db 758 ctgtgttgatatttgaacgaataaaacccatacaagaagctcccaagctcaaaa 817
 QY 61 AAAAAATGCAATCTCCGAGACAGCATGGAGATCGTATTCGAGACTTTAAGAAA 120
 Db 818 taaagtgaatttagaagatcatgagagtgatagaatctcgagatttaaaaa 877
 QY 121 CCAATCTGATGACAAACAGCAGCAACCAATCCACTTCTCATTTGCTGACATTGGACAAA 180
 Db 878 ccaataccatgagtgacacccactccactgctctcattcattgctgaagatcagtcacaa 937
 QY 181 GAATTGTGTAGTTCCTTGACAAATCTGGAAGCATGG 219
 Db 938 gaattgtgtcttagtcttcgtatagaatcgtgaagcatgg 976
 RESULT 15
 AAF92092
 ID AAF92092 standard; cDNA; 3265 BP.
 XX
 AC AAF92092;
 XX
 DT 15-MAY-2001 (first entry)
 XX
 DE Human PRO1124 cDNA.

XX
 KW Human; PRO protein; mapping; ss.
 XX
 OS Homo sapiens.
 XX WC200116318-A2.
 XX
 PD 08-MAR-2001.
 XX
 XX 24-AUG-2000; 2000WO-US23328.
 EF
 XX 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 07-DEC-1999; 99US-0169495.
 PR 09-DEC-1999; 99US-0170262.
 PR 13-JAN-2000; 2000US-0175481.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 03-MAR-2000; 2000US-0187202.
 PR 25-APR-2000; 2000US-0199397.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 05-JUN-2000; 2000US-0209832.
 XX
 XX (GETH) GENENTECH INC.
 XX
 XX Eason DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
 XX WPI: 2001-183260/18.
 DR P-PSDB: AAB87560.
 DR
 XX Eighty four nucleic acids encoding PRO polypeptides, useful in
 PT molecular biology, including use as hybridization probes, and in
 PT chromosome and gene mapping.
 XX
 XX Claim 2; Fig 69; 278pp; English.
 PS
 CC The present sequence for a human PRO polypeptide
 CC (secreted and transmembrane). The PRO protein, and PRO agonists, PRO
 CC antagonists or anti-PRO antibodies are useful for preparation of a
 CC medication useful in the treatment of a condition which is responsive to
 CC the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
 CC protein may also be employed as molecular weight markers for protein
 CC electrophoresis. The PRO coding sequence has applications in molecular
 CC biology, including use as hybridization probes, and in chromosome and
 CC gene mapping.
 XX
 XX Sequence 3265 BP; 1159 A; 596 C; 632 G; 878 T; 0 other;
 Query Match 60.3%; Score 132.6; DB 22; Length 3265;
 Best Local Similarity 75.3%; Pred. No. 4.9e-30;
 Matches 165; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 1 CTATAGTTGATTCGTGACAGCAACAAAGAGAGCTCCAAACAGCAAGCAATC 60
 Db 758 ctgtgttgatatttgaacgaataaaacccatacaagaagctcccaagctcaaaa 817
 QY 61 AAAAAATGCAATCTCCGAGACAGCATGGAGATCGTATTCGAGACTTTAAGAAA 120
 Db 818 taaagtgaatttagaagatcatgagagtgatagaatctcgagatttaaaaa 877
 QY 121 CCAATCTGATGACAAACAGCAGCAACCAATCCACTTCTCATTTGCTGACATTGGACAAA 180
 Db 878 ccaataccatgagtgacacccactccactgctctcattcattgctgaagatcagtcacaa 937
 QY 181 GAATTGTGTAGTTCCTTGACAAATCTGGAAGCATGG 219
 Db 938 gaattgtgtcttagtcttcgtatagaatcgtgaagcatgg 976

Thu Apr 4 09:27:45 2002

Search completed: April 4, 2002, 03:17:29
Job time: 60201 sec

us-09-049-696-5.rng

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 02:57:29 ; Search time 22700.8 Seconds

(without alignments)
159,879 Million cell updates/sec

Title: US-09-049-696-5

Perfect score: 220
Sequence: 1 CTATAGTTGATTCTGTACA.....GACAAATCTGGAAGCATGGC 220

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: 1: gb_da: 2: gb_hlg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_da: 16: em_fun: 17: em_hum: 18: em_in: 19: em_om: 20: em_or: 21: em_ov: 22: em_pat: 23: em_ph: 24: em_pl: 25: em_ro: 26: em_sts: 27: em_sy: 28: em_un: 29: em_vl: 30: em_hgo_hum: 31: em_hgo_inv: 32: em_hgo_rod: 33: em_hgo_hum: 34: em_hlg_inv: 35: em_hlg_rod: 36: em_hlg_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	220	100.0	2826	9	AF127036	AF127036 Homo sapi
2	220	100.0	3311	6	AF193489	AX139400 Sequence
3	220	100.0	3311	6	AF039400	AF039400 Homo sapi
4	218	99.1	35278	9	AF039401	AF039401 Homo sapi
5	218	99.1	113764	9	HS7651E10	AL122002 Human DNA
6	218	99.1	164891	2	AL358950	AL358950 Homo sapi
7	181.6	82.5	3079	4	AF095584	AF095584 Sus scrofa
8	166	75.5	2022	9	AK024970	AK024970 Homo sapi
9	162	73.6	2937	10	AB017156	AB017156 Mus muscu
10	132.6	60.3	3204	9	AF127035	AF127035 Homo sapi
11	132.6	60.3	3221	9	AK000072	AK000072 Homo sapi
12	132.6	60.3	3265	6	AX092338	AX092338 Sequence
13	131.2	59.6	140718	2	AL356270	AL356270 Homo sapi
14	103.2	46.9	1265	4	AF001264	AF001264 Bos tauru
15	103.2	46.9	2820	4	AF001263	AF001263 Bos tauru
16	103.2	46.9	3388	4	AF001262	AF001262 Bos tauru
17	103.2	46.9	3317	4	AF001261	AF001261 Bos tauru
18	102.2	46.5	2765	10	AF115852	AF115852 Mus muscu
19	102.2	46.5	3058	10	AF108501	AF108501 Mus muscu
20	102.2	46.5	3137	10	BC008147	BC008147 Mus muscu
21	99.4	45.2	3415	9	AF043976	AF043976 Homo sapi
22	99	45.0	3022	10	AF047838	AF047838 Mus muscu
23	99	45.0	3471	10	AF052746	AF052746 Mus muscu
24	98.8	44.9	175591	2	AC068071	AC068071 Homo sapi
25	96.8	44.0	2984	4	BT036445	U36445 Bos tauru
26	82.4	37.5	2832	9	AF127980	AF127980 Homo sapi
27	82.4	37.5	2970	9	AF043977	AF043977 Homo sapi
28	82.4	37.5	3604	9	AB026833	AB026833 Homo sapi
29	82.4	37.5	4077	6	AX054697	AX054697 Sequence
30	81	36.8	124096	6	AL354988	AL354988 Homo sapi
31	49.2	22.4	307	6	AX187657	AX187657 Sequence
32	49.2	22.4	333	6	AX186584	AX186584 Sequence
33	49.2	22.4	423	6	AX184779	AX184779 Sequence
34	39.8	18.1	182885	2	AC021760	AC021760 Homo sapi
35	38.2	17.4	120856	10	AC006944	AC006944 Mus muscu
36	38.2	17.4	220020	10	AC008100	AC008100 Mus muscu
37	38	17.3	173364	2	AC009807	AC009807 Homo sapi
38	38	17.3	173786	2	AP003069	AP003069 Homo sapi
39	37.8	17.2	127078	9	AP000708	AP000708 Homo sapi
40	37.6	17.1	95397	9	AL359632	AL359632 Human DNA
41	37.2	16.9	162973	2	AC016307	AC016307 Homo sapi
42	37.2	16.9	179056	2	AC078866	AC078866 Homo sapi
43	37.2	16.9	201195	2	AC068227	AC068227 Homo sapi
44	37	16.8	40662	3	CEY328128	AL031632 Caenorhab
45	37	16.8	74356	2	AC016511	AC016511 Homo sapi

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED
1	AF127036	Homo sapiens calcium-activated chloride channel protein 1 (CaCC1)	AF127036	AF127036.1	GI:4585468	human	Homo sapiens	1 (bases 1 to 2826)	Aguel, M., Vernat, T. and Culouscou, J.M.	Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea	FEBS Lett. 455 (3), 295-301 (1999)	99364503	10437792

REFERENCE 2 (bases 1 to 2826)
 AUTHORS Agnel, M. and Culouscou, J.-M.
 TITLE Direct Submission
 JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des Carrières, Rueil-Malmaison 92500, France

FEATURES
 Source Location/Qualifiers
 1. 2826
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="small intestine; colon"
 1. 2826
 /gene="CACCI"
 5. 2749
 /note="bovine epithelial chloride channel homolog"
 /codon_start=1
 /product="calcium-activated chloride channel protein 1"
 /protein_id="AA025487.1"
 /db_xref="GI:4385469"
 /translation="MGPKSSVFLLHLLEGALNSLIQLNNNGEGIVAVADPNVP
 EDELIOQIDMYTOASLYLEFATGKRFYKNAIILPEWTKADYVRLPEYKNA
 DVALESTPGNDPEYFEOGNGEGERHLPDFIAGKKLAEGPOGRAFEYMAH
 LRMGVEYNNDEKFEYLSNGRIQAVRCASGITGNVYKKGCGSCYTRKCFENVTGL
 YKCGEPVLOSROTERKASIMFAOHVDSIVFCPTQNNKAPKONKCNIRSTWEYL
 RSEDEKKTTPMTOPNPPTFSLQISQRTVCLVDNCSGSHATGNRLNLNQAGQLFL
 LQVEIGSWGWTPEDSAHVQSELQINGSDPDTLAKRLPAASGCTISCSGLRSA
 FTVIRKPYDGESEIVLTDGEDNTISGCFENEYSQSAIHTVALGSAQDELELSK
 MTGLOIYVADQVONNGLIDAFGLSSGNASVORSIOLESGLTLGNSOMNTEVY
 DSTYGRDTELTMTTQPOILLMDPSGOGGVYDKNTKMYLQIPIGAKVETWEL
 SIQASSOTLTITVYSRASNATPIPTITSTKNDTSKPEPSPLVYVYANIRGASPIILA
 SVTLISVNGKTVTLELDNGAGADTKDGYTSKFTYDINGRISVAKRALGVAS
 AARRVLPDQSGALYIPGWIENDEIQNNPPEPKNDVOHOKGCEFTSSGSGFVAS
 DVNAPIPDLEPFGQITDLKAEIHGSLMLNTWAPDDYDHGAHRYIIRISTSID
 LRDKNESLQVNTALIPKENSEVEVLEFPENTFENGDLFAIQAIVDKLSEI
 SNIAVSLFIPQIPETPSDETSAFCPNHINSTITGIIHKIMKMGELQSLIA"

BASE COUNT 875 a 623 c 632 g 696 t
 ORIGIN

Query Match 100.0%; Score 220; DB 9; Length 2826;
 Best Local Similarity 100.0%; Pred. No. 9 6e-50;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTATAGTTGAATTCGTGACAGACAAAACACAAAGAAAGCTCCAAACAGCAAAATC 60
 |||||
 DB 738 CTATAGTTGAATTCGTGACAGACAAAACACAAAGAAAGCTCCAAACAGCAAAATC 797
 |||||

QY 61 AAAAATGCAATCTCGAAGACATGGAAGATCCGATTCGACAGACCTTAAGAAAA 120
 |||||
 DB 798 AAAAATGCAATCTCGAAGACATGGAAGATCCGATTCGACAGACCTTAAGAAAA 857
 |||||

QY 121 CCATCCATATGACAAACAGACACCAAAATCCACCTTCTCATTCGTGAGATGGACAAA 180
 |||||
 DB 858 CCATCCATATGACAAACAGACACCAAAATCCACCTTCTCATTCGTGAGATGGACAAA 917
 |||||

QY 181 GAATTGTGTGTTAGTCTTGACAAATCTGGAAGCATGGC 220
 |||||
 DB 918 GAATTGTGTGTTAGTCTTGACAAATCTGGAAGCATGGC 957
 |||||

RESULT 2
 AX193489 3311 bp DNA PAT 15-AUG-2001
 LOCUS Sequence 1056 from Patent WO0149716.
 DEFINITION AX193489
 ACCESSION AX193489
 VERSION AX193489.1 GI:15211440
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3311)

AUTHORS Xu, J., Lodes, M. J., Secrist, H., Benson, D. R., Meagher, M. J.,
 Stolk, J. A., King, G. E., Wang, T., and Jiang, Y.
 TITLE Compounds for immunotherapy and diagnosis of colon cancer and
 JOURNAL methods for their use
 Patents: WO 0149716-A 1056 12-JUL-2001;
 CORIXA CORPORATION (US)

FEATURES
 Source Location/Qualifiers
 1. 3311
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="1p22-p31"
 /tissue_type="small intestine"
 1. 3311
 /gene="hCLCA1"
 352. 3096
 /note="transmembrane glycoprotein"
 /codon_start=1

BASE COUNT 1028 a 692 c 742 g 849 t
 ORIGIN

Query Match 100.0%; Score 220; DB 6; Length 3311;
 Best Local Similarity 100.0%; Pred. No. 9 6e-50;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTATAGTTGAATTCGTGACAGACAAAACACAAAGAAAGCTCCAAACAGCAAAATC 60
 |||||
 DB 1085 CTATAGTTGAATTCGTGACAGACAAAACACAAAGAAAGCTCCAAACAGCAAAATC 1144
 |||||

QY 61 AAAAATGCAATCTCGAAGACATGGAAGATCCGATTCGACAGACCTTAAGAAAA 120
 |||||
 DB 1145 AAAAATGCAATCTCGAAGACATGGAAGATCCGATTCGACAGACCTTAAGAAAA 1204
 |||||

QY 121 CCATCCATATGACAAACAGACACCAAAATCCACCTTCTCATTCGTGAGATGGACAAA 180
 |||||
 DB 1205 CCATCCATATGACAAACAGACACCAAAATCCACCTTCTCATTCGTGAGATGGACAAA 1264
 |||||

QY 181 GAATTGTGTGTTAGTCTTGACAAATCTGGAAGCATGGC 220
 |||||
 DB 1265 GAATTGTGTGTTAGTCTTGACAAATCTGGAAGCATGGC 1304
 |||||

RESULT 3
 AF039400 3311 bp mRNA PRI 14-DEC-1998
 LOCUS Homo sapiens calcium-dependent chloride channel-1 (hCLCA1) mRNA,
 DEFINITION complete cds.
 ACCESSION AF039400
 VERSION AF039400.1 GI:4009457
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3311)
 AUTHORS Gruber, A. D., Eblin, R. C., Ji, H. L., Schreier, K. D., Fuller, C. M. and
 Pauli, B. O.
 TITLE Genomic cloning, molecular characterization, and functional
 JOURNAL analysis of human CLCA1, the first human member of the family of
 MEDLINE Ca2+-activated Cl- channel proteins
 REFERENCE Genomics 54 (2), 200-214 (1998)
 99047526
 2 (bases 1 to 3311)
 AUTHORS Gruber, A. D., Eblin, R. and Pauli, B. O.
 TITLE Direct Submission
 JOURNAL Submitted (19-DEC-1997) Pathology, College of Veterinary Medicine,
 Cornell University, Ithaca, NY 14853, USA

FEATURES
 Source Location/Qualifiers
 1. 3311
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p22-p31"
 /tissue_type="small intestine"
 1. 3311
 /gene="hCLCA1"
 352. 3096
 /note="transmembrane glycoprotein"
 /codon_start=1

BASE COUNT 1028 a 692 c 742 g 849 t

ORIGIN

```

/product="calcium-dependent chloride channel-1"
/protein_id="AAC95428.1"
/db_xref="GI:4009458"
/translation="MGPFRKSSVFLLIHLLEGALNSGLQLNNGYEIGIVVADPNVPE
EDELIDQIKDMVTOASLYLFEATGRKFRFEYFNVALILPETKTRADYVIRPLEEYKRNA
DLVAESSTPGENDPEYFTEOMGNCCEKSERILTPDFLAGKLAIEGPOGAKFVEYMA
LRMVEFDEPNNDKEKFEYLSNGILOAVRSAGITGYNVAKCGGSGCYTRCGCFNFWTGL
YKCECFVLOSRORFEKASIMPAOHVDSITVECPTEONINKRANPKNOKCNLRISWEYI
RDSDEFKKTPTMTOPNPPTFSLQIGORVCIYLDLDSGSMATNSRLNRLNQAQLELT
LQYVELGSWGMVTFDSAAHVOSELILQINSGDDTDLAKRLPAASGATSLCSGLRSNA
VYFIRKRYPTDQSEIVLITDQEDNTIGSCFENGVQSQALITVYALGSAQDELELSK
MTGGLQYASDOYVONSGILDAFGALSSGANGVSORSILOESKGLTLNQQMMNTVYIL
DSYVGKDDFLITWTQPOLIMDPSGOKGGGVNPKRMVALIOLIGLAKVETMYK
SLOASQOTPLFVTSRASNATLPIRYTYSKNKTPSKVPSLYVYANIRQOASPIILRA
SVYALIESVNGKTYVTEILDNGACADATKDDGVYSRFTTIDIGRSVAKRALGGVNA
AARRVYIPQOOGALYIPGWLINENDIYQNNPPEYLPINKDOVHKQVCSRTSSGGSFVAS
DVAPNAPIDPLEPGCOLITDKAEIHGSSLINTLTWAPDGDYDHGHANHYIIRISYIIDL
LRDFNSLSIAVQNTALIPKKAENSEVELEFKENTTFENGSGFLPAIDAVDPKDSLESL
SNIRAVSLFIPQTPPEPSPEDTSAPCNPNIHNSITLPGIHLIKIMMKWIGELQSLSTA
"
```



```

exon      /gene="hclcal"
           /number=10
           /number=10
           /gene="hclcal"
           /number=11
           /number=11
           /gene="hclcal"
           /number=12
           /gene="hclcal"
           /number=13
           /number=13
           /gene="hclcal"
           /number=14
           /number=14
           /gene="hclcal"
           /number=15
           /number=15
exon
exon
exon
BASE COUNT 11245 a 6870 c 7230 g 9933 t
ORIGIN

```

```

Query Match 99.1%; Score 218; DB 9; Length 35278;
Best Local Similarity 100.0%; Pred. No. 3.7e-49;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 ATAGTGAATTCGTGTACAGAACAAACACACAAAGAGCTCCAAACAGCAAAATCAA 62
    |||||||
DB 18591 ATAGTGAATTCGTGTACAGAACAAACACACAAAGAGCTCCAAACAGCAAAATCAA 18650
    |||||||
QY 63 AAATGCAATTCGGAACACATGGAGTATCCGATTCCTGAGCACTTAAAGAAAACC 122
    |||||||
DB 18651 AAATGCAATTCGGAACACATGGAGTATCCGATTCCTGAGCACTTAAAGAAAACC 18710
    |||||||
QY 123 ACTGCTATGACACACACAGCCCAATCCACCTTCATTCGTCGAGATTGACAAAGA 182
    |||||||
DB 18711 ACTGCTATGACACACAGCCCAATCCACCTTCATTCGTCGAGATTGACAAAGA 18770
    |||||||
QY 183 ATTGTGTTAGTCCCTTGACAAATCTGGAAGCATGGC 220
    |||||||
DB 18771 ATTGTGTTAGTCCCTTGACAAATCTGGAAGCATGGC 18808
    |||||||

```

```

RESULT 5
HSJ651E10 113764 bp DNA PRI 27-MAY-2000
LOCUS Human DNA sequence from clone RP4-651E10 on chromosome 1p22.3-31.1,
DEFINITION complete sequence.
ACCESSION AL122002
VERSION AL122002.16 GI:8247274
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 113764)
Brown, A.
Direct Submission
Submitted (27-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Jun 4, 2000 this sequence version replaced gi:8247030.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

```

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>
 RP4-651E10 is from the library RPCI-4 constructed at the Roswell Park Cancer Institute by the group of Pleter de Jong. For further details see <http://bacpac.med.buffalo.edu/>
 VECTOR: pCYPAC2
 This sequence is the entire insert of clone RP4-651E10.

```

FEATURES
    source
        1. 113764
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /map="p22.3-31.1"
            /clone="RP4-651E10"
            /clone_11b="RPCI-4"
BASE COUNT 34890 a 21989 c 22351 g 34534 t
ORIGIN

```

```

Query Match 99.1%; Score 218; DB 9; Length 113764;
Best Local Similarity 100.0%; Pred. No. 4e-49;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 ATAGTGAATTCGTGTACAGAACAAACACACAAAGAGCTCCAAACAGCAAAATCAA 62
    |||||||
DB 18313 ATAGTGAATTCGTGTACAGAACAAACACACAAAGAGCTCCAAACAGCAAAATCAA 18372
    |||||||
QY 63 AAATGCAATTCGGAACACATGGAGTATCCGATTCCTGAGCACTTAAAGAAAACC 122
    |||||||
DB 18373 AAATGCAATTCGGAACACATGGAGTATCCGATTCCTGAGCACTTAAAGAAAACC 18432
    |||||||
QY 123 ACTGCTATGACACACAGCCCAATCCACCTTCATTCGTCGAGATTGACAAAGA 182
    |||||||
DB 18433 ACTGCTATGACACACAGCCCAATCCACCTTCATTCGTCGAGATTGACAAAGA 18492
    |||||||
QY 183 ATTGTGTTAGTCCCTTGACAAATCTGGAAGCATGGC 220
    |||||||
DB 18493 ATTGTGTTAGTCCCTTGACAAATCTGGAAGCATGGC 18530
    |||||||

```

```

RESULT 6
AL358950 164891 bp DNA HTG 23-JAN-2001
LOCUS Homo sapiens chromosome 1 clone RP11-444C12, *** SEQUENCING IN
DEFINITION PROGRESS ***
ACCESSION AL358950
VERSION AL358950.4 GI:12539689
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 164891)
Plumb, B.
Direct Submission
Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Jan 26, 2001 this sequence version replaced gi:998471.

```

COMMENT:
 ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquerry@sanger.ac.uk
 ----- Project Information
 Center project name: BA444C12

/protein_id="BAJ3743.1"
/db_xref="GI:3721912"
/translation="MESIKSEVFLIIHLLEGVLESLIQLNNNGEIVIAIDHDV
EDEALIIQIKDMVTOASPLIEATGKRFYFNVALIIPESKAPREYRPLEFKNA
DLVSTSPLENDPEYHIGAGEKGRIRHLPDPLAKRKTQYGGDDRFVEMAH
PRMGVFNENDEKFFYLSKGRQAVRCSAATGKNQVRCCGSCITNGKVIDRVTG
LYKNCVFPDPHONKASIMFNONINSVVECTEKNINDEAPNDQNRCMIRSTWEY
IOESDEKQVTPMTAOPAPPEFLIOLQRTVCVLKSGSMLDNDRLNNQASRLF
LIAJVEGSGSWGATPDSAAVVOSELKOLNAGADRLIKHLPYSAGTSCGLRT
AFVIKKKYPDGSSEIYLTDEEDNTISSCDLKOSGAIHTHTVALGPAAKELESL
KMTGLIOTSSDDOYNNGLVDAFALSNGMAIIOHSTOLERCGVINONNMMSVY
VDSVSGKDLFLITHTHTPTPIITMDISGVKNFILDITTKVAYLOVPGTAKGVK
YSDASSOTILITVSRASATLPPIITVNVNKNVGFSPVTVYASIRQASDILR
ASVVALLESVNGKVTLELDNGAGADAKNDVYSRPTAFDANGRYSVKIMAGVY
TSRORAAPPRNAMYIDGWIEDGEVRMNPRTSYQDKLCFSRTSGSFVATN
VPAAPAPIDLEPCQITDLKASIOGNILNLTAPGDYDHGRASNTIIMRSTSYD
LRHENTSLQVNTGLIPKASSSEIFEFEFGNGTDIFALIOAVDMSNKSEI
SNIRASVFIPADPPIEDSTPPCPDISINSTIPGILHVKIMKMLGEMQVTLGLH"
BASE COUNT 860 a 718 c 693 g 666 t
ORIGIN
Query Match 73.6%; Score 162; DB 10; Length 2937;
Best Local Similarity 83.9%; Pred. No. 5.6e-34;
Matches 183; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 1 CTATAGTTGAATTCGTGTACAGAACAAACCAAGAGCTCCAAACAGCAAAATC 60
Db 751 CTGTGGTGAATTCGTGTACAGAACAAACCAAGAGCTCCAAACAGCAAAATC 810
QY 61 AAAATGCAATTCGGAAGACAGATGAGATGCGTATTCGTGAGACTTAAAGAAA 120
Db 811 AAGATGCAATTCGGAAGACAGATGAGATGCGTATTCGTGAGACTTAAAGAAA 870
QY 121 CCATCCCTATACAAACAGACCAATCCCACTTCTCATGCTGCAAGATGAGAAA 180
Db 871 CCATCCCTATACAAACAGACCAATCCCACTTCTCATGCTGCAAGATGAGAAA 930
QY 181 GAATGTGTGTAGTCTTGAACAAATCTGGAAGCATG 218
Db 931 GAATGTGTGTAGTCTTGAACAAATCTGGAAGCATG 968
RESULT 10
AF127035 3204 bp mRNA PRI 11-AUG-1999
LOCUS Homo sapiens calcium-activated chloride channel protein 2 (CACCC2)
DEFINITION mRNA, complete cds.
ACCESSION AF127035
VERSION AF127035.1 GI:5726288
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3204)
AUTHORS Agnel,M., Vermaat,T. and Culouscou,J.M.
TITLE Identification of three novel members of the calcium-dependent
chloride channel (CACCC) family predominantly expressed in the
digestive tract and trachea
JOURNAL FEBS Lett. 455 (3), 295-301 (1999)
MEDLINE 99364503
PUBMED 10437792
REFERENCE 2 (bases 1 to 3204)
AUTHORS Agnel,M. and Culouscou,J.-M.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-1999) Groupe Genomique, SYNTHELABO, 10 Rue des
Carrières, Rueil-Malmaison 92500, France
FEATURES
source 1..3204
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="colon"

gene
1..3204
/gene="CACCC2"
CDS 29..2782
/gene="CACCC2"
/note="bovine epithelial chloride channel homolog"
/codon_start=1
/product="calcium-activated chloride channel protein 2"
/protein_id="AAD48398.1"
/db_xref="GI:5726289"
/translation="MGLEPRGVEFLVLCILHOSNPSFKLNNNGEDVIVYDVSPE
DEKITEQJEDWATTAATSTLFEATERFRFFKANSIIPENKENDPQKPKENHKNAD
VIVAPPTLPGRDEPTKQPTREGGEGETHPTPLDLKKNQNGPPKLPVHEVAHL
RMGVEDEVNEQDPEFRASKIEARSGAGISGRNRYKCOGSCLSRACRLDYSTTKL
YKDCQDFPNDQVQTEKASIMFQSDISVVEECNEKTHQEPASLONICNPSREVEI
SNSSEFKTIPMPVPPPPVSLKIRORICVLYDKSGMGRKLRNMOAAHFL
LOIYENSGWGMVHEDSRATYINKLIQKSSDERNTLAAGLPYTLGCTSCGKRYA
FOYIGELISQDSESVYLLTDEEDNTASSCIDERKQSAIVHFLNAGRADVAIEMS
KITGSHRYVDEAQNGLIDAFGLATSGNDLSQSLQSLSKGLTLLSNMAMNDTVI
IDSTVGKDFELITVNSLPPSISIMDPGCTIMENFTVAKNKNKDVNSPSPAIYVAEILQGVV
YNIQAKNPDELITVTSRAANSVSPITVAKNKNKDVNSPSPAIYVAEILQGVV
LGANTVATIESONGHEVLELDNGAGADSRKNDGVSRFPVATYENGSRYSIKVRAG
GANTARLIRPPIPNAAIYRGVYNGELEAPPREIDEDVOTIEDSRSSAGAV
VSQYPSLEPPOYPPSQITDDATVHEKILITWTAPCDNDVGKQVNTIRISAIL
DLRQSEDDALQVNTTDLSPKANSKESFAFENISEBNATHIFAISIDKSNLTSK
VSNIAQVTLFTPOANPDIDPTPTPTPPDKSHNSGVNISTLVLSIGSVIVNFILS
TTI"
BASE COUNT 1098 a 594 c 633 g 879 t
ORIGIN
Query Match 60.3%; Score 132.6; DB 9; Length 3204;
Best Local Similarity 75.3%; Pred. No. 5.5e-26;
Matches 165; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 1 CTATAGTTGAATTCGTGTACAGAACAAACCAAGAGCTCCAAACAGCAAAATC 60
Db 762 CTGTGGTGAATTCGTGTACAGAACAAACCAATACCAAGAGCTCCAAACAGCAAAATC 821
QY 81 AAAATGCAATTCGGAAGACAGATGAGATGCGTATTCGTGAGACTTAAAGAAA 120
Db 822 TAAAGTGCATTTTGAAGATACATGAGAGTGAATTCGTGAGACTTAAAGAAA 881
QY 822 TAAAGTGCATTTTGAAGATACATGAGAGTGAATTCGTGAGACTTAAAGAAA 881
QY 121 CCATCCCTATACAAACAGACCAATCCCACTTCTCATGCTGCAAGATGAGAAA 180
Db 882 CCATCCCTATACAAACAGACCAATCCCACTTCTCATGCTGCAAGATGAGAAA 941
QY 181 GAATGTGTGTAGTCTTGAACAAATCTGGAAGCATG 219
Db 942 GAATGTGTGTAGTCTTGAACAAATCTGGAAGCATG 980
RESULT 11
AK000072
LOCUS 3221 bp mRNA PRI 22-FEB-2000
DEFINITION Homo sapiens CDNA FLJ20065 fls, clone COL01613, highly similar to
ECIC BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN.
ACCESSION AK000072
VERSION AK000072.1 GI:7019922
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens colon CDNA to mRNA, clone, lib.COL clone:COL01613.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL unpublished (2000)
REFERENCE 2 (bases 1 to 3221)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission

JOURNAL	Submitted (15-FEB-2000) to the DDBJ/EMBL/GenBank databases. Sunio Supano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, TOKYO 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)					
COMMENT	NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; CDNA full insert sequencing. Research Association for Biotechnology; CDNA library construction. 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).					
FEATURES						
SOURCE	Location/Qualifiers 1..3221 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="COL01613" /clone_1lb="COL" /tissue_type="colon" /note="cloning vector pME18SF13"					
MISC_FEATURE	1..3221 /note="highly similar to ECIC_BOVIN EPITHELIAL CHLORIDE CHANNEL PROTEIN"					
BASE COUNT	1105 a	600 c	634 g	882 t		
ORIGIN						
Query Match	60.3%; Score 132.6; DB 9; Length 3221;					
Best Local Similarity	75.3%; Pred. No. 5,5e-26;					
Matches 165; Conservative	0; Mismatches 54; Indels 0; Gaps 0;					
OY	1	CATAGTTGAATTCTGTACAGAACAAACCACACAAGAAGACTCCMAAAGCAAAATC	60			
Dd	776	CGTGGTGAAATTTTTGTAACGAAAAACCCATAATCAAAACCTCCAAACCTCAAAAACA	835			
OY	61	AAAATGCATCTCCGAAGCACATGGGAAGTAGTCGCGTATCTCTGAGCATTTAAGAAA	120			
Dd	836	TAAATGCAATTTTGAAGTACATGGGAGGTATATAGCAATTCGAGATTTTAAAAACA	895			
OY	121	CCAGTCCTATGACAAACACAGCCACCAAAATCCCACCTTCATTTGCTGCAGATTGCACAA	180			
Dd	896	CCATACCCATGSGTGACACACCACTCCTCCACCTGTCTTCATTTGCTGAAGATCCGTCANA	955			
OY	161	GAAATGCTGTTAGTCCCTTGACAAATCTGGGAAGCATGG	219			
Dd	956	GAAATGTGTGCTTACTGTTGATATAGTCTGGAAGCATGG	994			
RESULT 12						
AX092338						
LOCUS	AX092338	3265 bp	DNA	PAT	21-MAR-2001	
DEFINITION	Sequence 69 from Patent WO0116318.					
ACCESSION	AX092338					
VERSION	AX092338.1 GI:13444483					
KEYWORDS	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 3265) Eaton,D.L., Flivaroff,E., Gerritsen,M.E., Goddard,A., Godowski,P.J., Grimaldi,C.J., Gurney,A.D., Watanabe,C.K. and Wood,W.I. Secreted and transmembrane polypeptides and nucleic acids encoding the same					
AUTHORS	Patent: WO 0116318-A 69 08-MAR-2001; Genentech, Inc. (US) Location/Qualifiers 1..3265 /organism="Homo sapiens" /db_xref="taxon:9606"					
BASE COUNT	1159 a	596 c	632 g	878 t		
ORIGIN						

```

Query Match Similarity      60.3%; Score 132.6; DB 6; Length 3265;
Best Local Similarity      75.3%; Pmed No.5.5e-26;
Matches 165; Conservative 0; Mismatch 54; Indels 0; Gaps 0;

QY 1 CTATGTTGAATTCCTGTGCAGAACAAACCCACACAAAGAGAGCTCCAAACAAACAAATC 60
Db 4738 CCGTTTGGAAATTTTGTATACGAAAAAACCCATATACAAAGAGCTCCAAAGCCACAAACA 817
QY 18 AAAATGCAATCTCCGAAGCACATGGGAAGTATCCGTGATTTCTGAGACATTTAAGAAAA 120
Db 618 TAAATGTGAATTTTGAAGATACATGGGAGGTGATTGACAAATTCGAGGATTTTAAAAACA 877
QY 1121 CCACCTCATGACACACAGCACCACAAATCCACCTCTCATGCTGCAGATGTGACAAA 180
Db 878 CCATACCCAGTGTGACACACACCTCTCCACCTGCTTCTCATGCTGGAAGATCACTCAA 937
QY 181 GAATGTGTGTTAGTCTCTGACAAATCTGGAAGCATGG 219
Db 598 GAATGTGTGCTTACTGCTTGTATAGTCTGGAAGCATGG 976

RESULT 13
AL356270 140718 bp DNA HTG 23-JAN-2001
LOCUS Homo sapiens chromosome 1 clone RP5-873015, *** SEQUENCING IN
PROGRESS ***, 22 unordered pieces.
AL356270
AL356270.3 GI:9797470
HTG, HTGS_PHASE1, HTGS_DRAFT.
KEYWORDS human:
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 140718)
AUTHORS Pavlitt.R.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
On Aug 12, 2000 this sequence version replaced gi:9213800.

COMMENT:
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
Project Information
Center project name: d1873015
----- Summary Statistics -----
Assembly program: XAPP, version 4.5
Sequencing vector: M13, M77815; 12% of reads
Sequencing vector: plasmid; 108752; 87% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 129686 bases at least Q40
Consensus quality: 134097 bases at least Q30
Consensus quality: 136445 bases at least Q20
Insert size: 138618; sum-of-contigs
Insert size: 144145; 1.2% error; agarose-fp
Quality coverage: 3.10x in Q20 bases; sum-of-contigs Quality
coverage: 3.21x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 3292: contig of 3292 bp in length
* 3293 3392: gap of 100 bp
* 3393 10168: contig of 6776 bp in length

```


AUTHORS Elble, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R., Goodwin, A., and Pauli, B.U.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-1997) Vet. Pathology, Cornell U., VMC C4-146, Ithaca, NY 14853, USA

FEATURES

source location/Qualifiers
 1. 1265
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /tissue_type="lung"
 /cell_type="aortic endothelial cells"
 /clone="4"
 109. 1137
 /note="endothelial adhesion molecule; chloride channel; secreted form"
 /product="Lu-ECAM-1"
 /protein_id="AAB86532.1"
 /db_xref="GI:2623769"
 /translation="MYCLNVLFLHLPGKSSMVNLINNGYDGIYAIINPSVE
 DEKLINIKEMYTEASTYLEPHATKRVYFRNYSILIPMTKSEYFIRKESYDAD
 VIVANPYLYGDDPYLYOGRGEGKGYIHPNPLLTNNFHYSGRHFVEMAH
 RMGIDDEVNDOPFYISRNNTIATRCSTHTINGINVFYKCGSCITSLCRDSOTG
 LYBAKCTFLPKRSQTAKESSIMPSLSHTEFECTHTNTEAPNLONKMCNGKSTWDY
 INNSVDFQNTSPMTENMPHTPHFSLKSKORVCLVDKSGMSAEDITYLALLIKI
 FLIGNNTI"

BASE COUNT 440 a 231 c 231 g 363 t

Query Match 46.9%; Score 103.2; DB 4; Length 1265;
 Best Local Similarity 69.5%; Pred. No. 5.2e-18;
 Matches 157; Conservative 0; Mismatches 63; Indels 6; Gaps 1;

QY 1 CTATAGTGAATTCCTGACAGCAAAACCAACCAAGAGCTCCAAAGCAAAATC 60
 || | ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 845 CTGTGACTGATTTTGTACGAGAAAACACACATACAGAGCTCCAAACCTACAAACA 904
 QY 61 AAAATGCAATCTCCAGACACATGGAGTGCATGATTCGTGAGACTTTAAGAAA 120
 || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 905 AAATGTCATGGCAAAAGACATGGATGTATGATGACTCTGTTGATTTCAAGATA 964
 QY 121 CCAGTCTTATGACA-----ACACAGCCACCAAAATCCACCTTCTCATTCCTGCAGATTG 174
 | ||||| ||||| || | || | ||||| ||||| ||||| ||||| ||||| |||||
 DB 965 CACTCTCCATGACAGAAATGATCAACCGACTCATCTTCAATTTTCATGCTCAAGTCA 1024
 QY 175 GACAAAGAAATGTGTGTAGTCTTGACAAATCTGGAAGACATGGC 220
 || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1025 AACAGCGGAGTAGTCTGTGTGTTGATTAATCTGGAAGCATGTC 1070

RESULT 15
 LOCUS AF001263 2820 bp. mRNA MAM 18-NOV-1997
 DEFINITION Bos taurus clone 3 endothelial adhesion molecule Lu-ECAM-1 mRNA,
 complete cds.
 ACCESSION AF001263
 VERSION AF001263.1 GI:2623766
 KEYWORDS
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 2820)
 AUTHORS Elble, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
 Goodwin, A., and Pauli, B.U.
 TITLE Cloning and characterization of Lu-ECAM-1 suggest it is an
 endothelial chloride channel
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2820)
 AUTHORS Elble, R.C., Widom, J., Gruber, A.D., Abdel-Ghany, M., Levine, R.,
 Goodwin, A., and Pauli, B.U.

TITLE Direct Submission
 JOURNAL Submitted (25-APR-1997) Vet. Pathology, Cornell U., VMC C4-146, Ithaca, NY 14853, USA

FEATURES

source location/Qualifiers
 1. 2820
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone="3"
 /tissue_type="lung"
 /cell_type="aortic endothelial cells"
 194. 2656
 /note="endothelial adhesion molecule; chloride channel"
 /product="Lu-ECAM-1"
 /protein_id="AAB86531.1"
 /db_xref="GI:2623767"

BASE COUNT 940 a 565 c 560 g 755 t

Query Match 46.9%; Score 103.2; DB 4; Length 2820;
 Best Local Similarity 69.5%; Pred. No. 5.4e-18;
 Matches 157; Conservative 0; Mismatches 63; Indels 6; Gaps 1;

QY 1 CTATAGTGAATTCCTGACAGCAAAACCAACCAAGAGCTCCAAAGCAAAATC 60
 || | ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 930 CTGTGACTGATTTTGTACGAGAAAACACACATACAGAGCTCCAAACCTACAAACA 989
 QY 61 AAAATGCAATCTCCAGACACATGGAGTGCATGATTCGTGAGACTTTAAGAAA 120
 || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 930 AAATGTCATGGCAAAAGACATGGATGTATGATGACTCTGTTGATTTCAAGATA 1049
 QY 121 CCAGTCTTATGACA-----ACACAGCCACCAAAATCCACCTTCTCATTCCTGCAGATTG 174
 | ||||| ||||| || | || | ||||| ||||| ||||| ||||| ||||| |||||
 DB 1050 CACTCTCCATGACAGAAATGATCAACCGACTCATCTTCAATTTTCATGCTCAAGTCA 1109
 QY 175 GACAAAGAAATGTGTGTAGTCTTGACAAATCTGGAAGACATGGC 220
 || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1110 AACAGCGGAGTAGTCTGTGTGTTGATTAATCTGGAAGCATGTC 1155

Search completed: April 4, 2002, 02:58:21
 Job time: 60183 sec